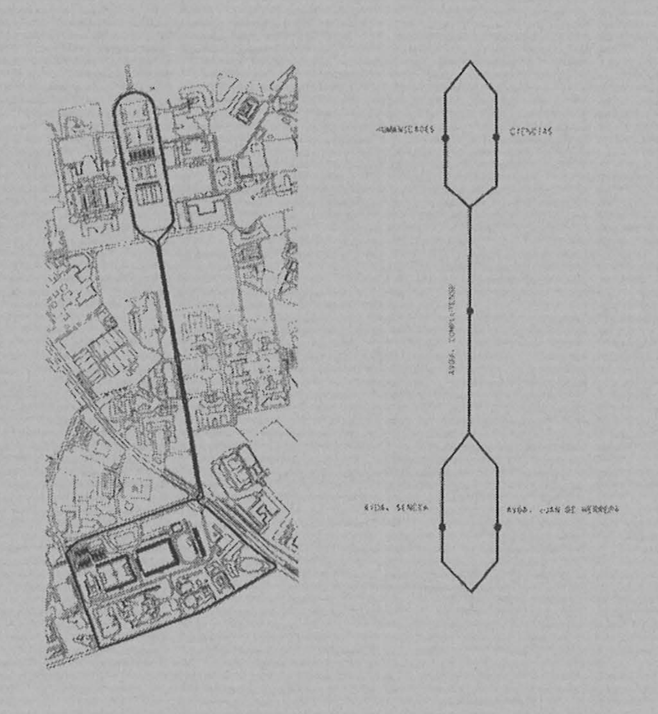
LA IMAGEN DE LOS FENÓMENOS GEOGRÁFICOS EN UN SISTEMA DE INFORMACIÓN GEOGRÁFICA

por

PILAR CHÍAS NAVARRO



CUADERNOS

DEL INSTITUTO

JUAN DE HERRERA

DE LA *ESCUELA DE*

ARQUITECTURA

DE *MADRID*

5-11-04

LA IMAGEN DE LOS FENÓMENOS GEOGRÁFICOS EN UN SISTEMA DE INFORMACIÓN GEOGRÁFICA

por

PILAR CHÍAS NAVARRO

CUADERNOS
DEL INSTITUTO
JUAN DE HERRERA
DE LA *ESCUELA DE*
ARQUITECTURA
DE MADRID

5-11-04

CUADERNOS
DEL INSTITUTO
JUAN DE HERRERA

0 VARIOS
1 ESTRUCTURAS
2 CONSTRUCCIÓN
3 FÍSICA Y MATEMÁTICAS
4 TEORÍA
5 GEOMETRÍA Y DIBUJO
6 PROYECTOS
7 URBANISMO
8 RESTAURACIÓN

NUEVA NUMERACIÓN

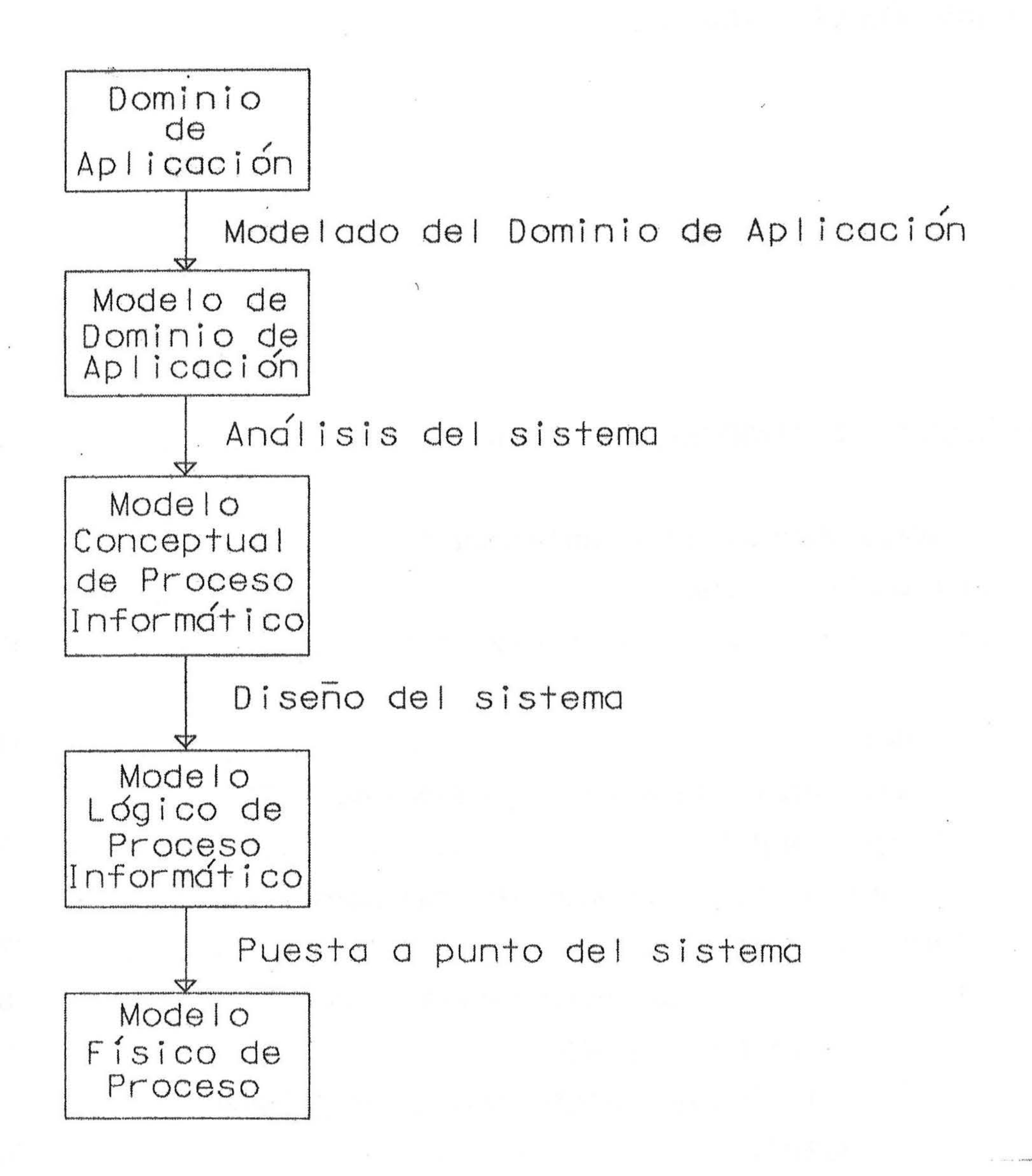
5 Área
11 Autor
04 Ordinal de cuaderno (del autor)

La imagen de los fenómenos geográficos en un sistema de información geográfica

Instituto Juan de Herrera.
Escuela Técnica Superior de Arquitectura de Madrid.
Gestión y portada: Nadezhda Vasileva Nicheva
CUADERNO 173.01 / 5-11-04
ISBN: 84-9728-116-0
Depósito Legal: M-44370-2004

LA IMAGEN DE LOS FENÓMENOS GEOGRÁFICOS EN UN SISTEMA DE INFORMACIÓN GEOGRÁFICA

Índice



Esquema del proceso de diseño de un SIG

LOS DATOS DEL DOMINIO DE APLICACIÓN

Antes de que se pudieran aplicar los ordenadores a la Cartografía, todos los tipos de mapas -topográficos, temáticos, cartas- tenían en común el hecho de que la base de datos espacial que contenían consistía en un dibujo sobre papel u otro soporte, que a lo sumo se completaba con unas tablas.

En ellos, la información se codificaba en forma de entidades básicas -puntos, líneas o superficies- combinadas para definir una serie de símbolos, que además se representaban utilizando una gran variedad de recursos gráficos tales como la codificación de los tipos y espesores de línea, el color o la rotulación, cuyos significados eran explicados en la *leyenda* del mapa.

El mapa era, por tanto, un documento cerrado, estático y correspondiente a un determinado momento histórico; pero además, era el resultado de la visión personal y de la aplicación de los criterios -de selección, simplificación y grafiación- de su autor.

Recientemente, al incorporarse las imágenes procedentes de aviones o de satélites, se han podido apreciar los cambios que tienen lugar sobre la superficie terrestre -evolución de la desertificación, de los núcleos de población, variaciones en los usos del suelo y en la cubierta vegetal, etc.-, obteniéndose secuencias que ofrecen visiones dinámicas de los fenómenos que sobre ella tienen lugar.

Pero los productos de la Teledetección suelen consistir en grandes cantidades de datos que se almacenan en cintas magnéticas o en otros dispositivos de almacenamiento de gran capacidad. Además, estos datos digitales no se presentan con las formas tradicionales de punto, línea y superficie, sino codificados en celdas o *pixels* en una matriz bidimensional que contienen simplemente un número que indica, por ejemplo, la cantidad de energía electromagnética reflejada por la superficie terrestre en una posición determinada sobre ella.

Como consecuencia de estas nuevas exigencias fue necesario diseñar nuevas "herramientas" para manipular tales datos numéricos, y transformarlos en imágenes en las que se pudieran reconocer los fenómenos de la parte de superficie terrestre correspondiente. Un paso más permitiría geo-referir las imágenes a la red geodésica para determinar las posiciones exactas de los objetos.

Y de este modo, y no exclusivamente de la mano de los cartógrafos, sino de los matemáticos, los físicos y los informáticos -frecuentemente al servicio de los ejércitos-, se fue gestando una nueva aproximación a la Cartografía, estrechamente vinculada a otros ámbitos tecnológicos y científicos como la Teledetección, la Informática, la Geodesia, las Ciencias de la Tierra, etc.

Paralelamente, en las décadas de los 60 y los 70 se gestaron nuevas tendencias en la utilización de los datos, al considerarse que los diferentes fenómenos que tienen lugar en la superficie terrestre están interrelacionados y no son independientes unos de otros. Arrancaron entonces los primeros proyectos integrados e interdisciplinares, que partieron de dos líneas de investigación diferentes.

La primera, encabezada por Hopkins y definida como el "Método Gestaltista", perseguía la definición de *unidades naturales* reconocibles, descriptibles y cartografiables en función de la total interacción de los atributos que se estudiaban; esta tendencia fue adoptada en la gestión de recursos por organismos tan importantes como el UK Land Resources Division (hoy el Land Resources Development Center del Reino Unido) o el ITC de Entschede (Holanda).

La segunda, descrita y utilizada por el arquitecto-paisajista estadounidense McHarg (1969) y por muchos urbanistas, se basaba en la combinación e integración de datos procedentes de fuentes especializadas variadas -por simple superposición de copias transparentes de los mapas-fuente sobre una mesa con luz-, y también en la observación de las coincidencias de contornos.

Fue otro arquitecto urbanista americano, que fuera luego director del Harvard Graduate School of Design's Laboratory for Computer Graphics, Howard T. Fisher quien desarrolló en 1963 la idea que había tenido E.M. Horwood de utilizar el ordenador para hacer mapas sencillos, imprimiendo líneas de valores estadísticos como una malla sobre papel blanco. Denominó a su programa SYMAP (*SYnagraphic MAPping system*, del griego *synagein*, reunir), y le incorporó unos módulos para el análisis de los datos que permitían su manipulación para obtener isolíneas por interpolación, a la vez que hacían posible su representación gráfica a base de sucesivas impresiones de caracteres para obtener diferentes tonos de gris.

El SYMAP fue el primero de una serie de programas de cartografiado automático como los también célebres programas *raster* -o de malla de celdas- GRID e IMGRID,

que permitían informatizar el método de superposición de McHarg. Además, los desarrolladores pronto descubrieron que, con muy poco esfuerzo adicional, podían incorporar análisis lógicos y espaciales de los datos con aplicación directa en planeamiento y estudio del medio ambiente, lo que hasta el momento se venía haciendo penosamente a mano, y con muchas limitaciones.

Y aunque las primeras salidas gráficas eran bastante deficientes y fueron rechazadas por muchos cartógrafos, a finales de los 70 los avances en el dise□o de los programas, en sus capacidades analíticas y de representación gráfica, y en la captura automática de los datos, habían sido tales que ya no permitían dudar de que la *Cartografía Asistida por Ordenador* (CAO) era la herramienta del futuro, siempre que la introducción de esta nueva tecnología se realizase con la mente lo suficientemente abierta para afrontar cambios en los hábitos de trabajo tradicionales.
Además, en paralelo habían evolucionado también los campos científicos asociados como la cartografía catastral, topográfica y temática; la ingeniería civil, la geografía, la matemática, la geología, la fotogrametría, el urbanismo, las redes de infraestructuras, la teledetección y la manipulación y análisis de imágenes.

El resultado de tantos esfuerzos desarrollados independientemente, ha sido la coexistencia hoy de un gran número de posibilidades de procesado de los datos espaciales, que pueden integrarse en lo que se ha denominado un Sistema de Información Geográfica o SIG.
Por tanto, un SIG es una tecnología compleja que permite introducir, manejar, analizar y presentar datos geográficos de procedencia variada, a partir de los cuales es posible obtener información sobre distribuciones y procesos espaciales y temporales que tienen lugar sobre la superficie terrestre.
Para ello, el sistema se compone de *hardware,* de *software* y de datos, y requiere de cierta formación especializada por parte de sus usuarios. La importancia de éstos nunca debe minimizarse, no sólo porque deben intervenir activamente en la puesta a punto del SIG, sino porque de ellos dependerá finalmente el éxito o el fracaso de su implantación.

1. Las etapas del modelado de un Sistema de Información Geográfica

Para que los datos relativos a los fenómenos geográficos puedan ser introducidos en un ordenador, el usuario debe estructurarlos de manera que puedan ser procesables, lo que no siempre supone organizarlos de un modo semejante a como requeriría una elaboración no informatizada. De hecho, la manera en que se perciben los fenómenos geográficos no es habitualmente el mejor modo de estructurar la base de datos.
Por otra parte, los datos han de grabarse y almacenarse en dispositivos magnéticos siguiendo un proceso que ha de estar organizado de un modo determinado.

El proceso puede resumirse en los pasos siguientes:

1º Percepción de la estructura del fenómeno por el usuario

⇕

2º Representación en el SIG de la estructura del fenómeno

⇕

3º Estructura de la Base de Datos

⇕

4º Estructura del *hardware*

Desde otro punto de vista complementario, el desarrollo de un SIG pasa por diferentes etapas que van desde el estudio de su viabilidad, su concepción y su diseño -que han de adecuarse a los fines perseguidos- hasta su puesta a punto y su desactivación[1], circunstancias que incluyen su reconsideración, mantenimiento y actualización a lo largo de toda su vida útil.
Este ciclo se inicia analizando para qué se quiere el sistema y definiendo los objetivos para los que se diseña, con el fin de satisfacerlos del modo más eficaz posible. Posteriormente, y durante todas las etapas del proceso de diseño e implementación se recomienda seguir un proceso iterativo y continuo en el que los supuestos que se hayan adoptado en un principio puedan ser revisados.

[1] Etapa en la que el SIG se deja en manos del usuario, con todas sus consecuencias.

Una vez fijados los objetivos, es necesario acotar el ámbito de estudio o *dominio de aplicación*, y definirlo conceptualmente en términos matemáticos utilizando un *modelo de información espacial* adecuado y procesable por medio del ordenador. Esta fase es independiente de cualquier tipo de proceso informático específico, y se centra en definir las características de todos las entidades, procesos o fenómenos que interesen y que se encuentren incluidos en el ámbito de estudio, así como todas las relaciones relevantes que entre ellos puedan establecerse.

El término *modelo* corresponde al concepto general de abstracción del mundo real, que extrae de él sólo las propiedades que resultan interesantes para una aplicación concreta. Es, por tanto, una construcción artificial en la que una parte de un *dominio fuente* -que es el dominio de aplicación- se representa en otro *dominio final.*

El propósito del modelo es simplificar y abstraer el dominio fuente permitiendo su visión y análisis en un nuevo contexto tal, que cualquier deducción, consecuencia o resultado que en el dominio final se obtenga pueda ser de aplicación en el dominio fuente.

Un ejemplo claro de modelo es un simulador de vuelo, que representa en un entorno artificial objetos del mundo real tales como la cabina de un avión con su instrumentación, sonidos, movimientos, vistas y, en definitiva, el espacio de navegación. Las experiencias con el simulador pueden ser trasladadas posteriormente a vuelos reales.

Otro ejemplo más próximo a la tecnología SIG es el que brinda la Cartografía, en la cual el mundo geográfico es el dominio fuente, que es modelado hasta obtener un mapa -dominio final. Un viajero que quiera ir de Madrid a El Escorial habrá de consultar el mapa y después tendrá que trasladar sus conclusiones al mundo real eligiendo un itinerario adecuado. Obviamente, cuanto mejor sea el mapa, mejor será su viaje.

Para que un modelo sea útil debe ajustarse lo más posible al dominio fuente, y debe permitir una comunicación fácil entre este dominio y el final. Estas propiedades definen matemáticamente una función de un dominio en otro, que mantiene parte de la estructura en la aplicación. En los capítulos siguientes se verán las posibilidades de modelado del mundo real que manejan habitualmente los SIG.

Por ejemplo, si se quisiera estudiar el abastecimiento eléctrico de la Ciudad Universitaria con el fin de prever líneas alternativas en caso de cortes de energía, el modelo adecuado se definiría matemáticamente como una red. En cambio, si se quisiera cartografiar la radiación solar sobre los planos inclinados que conforman los citados terrenos, y conociendo de éstos su posición geográfica -a efectos de la inclinación de los rayos solares a lo largo del año-, habría que definir un modelo matemático que permitiese calcular la suma de las radiaciones directa, reflejada y difusa en cada punto del ámbito de estudio.
Un *modelo de datos* en la terminología de los SIG, supone una concreción de la definición anterior, y consiste en un conjunto de directrices -los nombres de las unidades lógicas de los datos y las relaciones que entre ellos existen-, que permiten la representación lógica de los datos en una base de datos.

La segunda fase que debe abordarse una vez definido el *modelo del dominio de aplicación*, es la construcción de un *modelo conceptual informatizable* teniendo en cuenta el soporte informático de todo el sistema. Pues aunque el ojo humano es muy eficiente a la hora de reconocer formas, el ordenador necesita ser instruido exactamente en cómo manejar y mostrar las distribuciones espaciales.
En esencia existen dos maneras diferentes, y podría decirse que complementarias, de representar los datos espaciales en el ordenador, y que podrían calificarse respectivamente de *implícita* y *explícita*.

- La representación explícita o *raster* supone que la forma se construye como un conjunto de puntos en una malla, y para que el ordenador sepa qué es cada punto se le asigna un valor numérico, un color o un tono de gris.
- En cambio, la representación implícita utiliza un conjunto de puntos, líneas y superficies, que se definen a través de sus puntos y de algún tipo de conectividad. Por ejemplo, los puntos extremo e inicial de un segmento definirán vectores que representan la forma del objeto en cuestión; además, unos indicadores entre las líneas mostrarán al ordenador cómo se conectan los segmentos para conformar el objeto.

Para diseñar estos modelos se recurre a métodos como variados. Son ejemplos de ellos el llamado "Entidad-Relación" o el "modelado de Objetos", y suelen ser

desarrollados por especialistas. No obstante, su conocimiento por parte del usuario resulta fundamental a la hora de optimizar el sistema.

La tercera fase persigue diseñar el *modelo lógico de computación,* en la que partiendo del modelo del dominio de aplicación, se elige un *modelo de información espacial* que se integra dentro de un tipo de proceso informático concreto para mantener los datos. Por ejemplo, una posibilidad es integrar los datos no espaciales o atributivos en bases de datos relacionales, y representar los datos espaciales como esquemas de relación. Pero existen otras posibilidades como las que brindan las bases de datos orientadas a objetos, las deductivas, etc.

Aunque no es fundamental que el usuario conozca en detalle como se pueden organizar los datos dentro del ordenador, el comprender los diferentes métodos de estructuración de los datos le ayudará a saber cómo funciona el sistema, sus ventajas y limitaciones.

Por otra parte, las características más importantes que debe presentar una base de datos son:

- que debe permitir un rápido acceso a los mismos
- y que aquéllos puedan cruzarse con facilidad y en el mínimo de tiempo.

No existe un método que sea "mejor" y que se pueda recomendar para cualquier situación, y por ello son varios los que se utilizan y se presentan en el capítulo correspondiente. También por este motivo se continúa investigando en el ámbito de los *gestores de bases de datos*, que son los programas que controlan la entrada, salida, almacenamiento y recuperación de los datos de una base de datos.

Finalmente, el diseño del *modelo físico del proceso* supone la puesta a punto de los modelos anteriores sobre un sistema y plataforma específicos y, en definitiva, la representación del mundo real como un conjunto de fenómenos procesables por un ordenador.

En los SIG, los modelos operan sobre una gran variedad de situaciones, y pueden ser sencillos o de una gran complejidad.

1.1 La calidad del modelo: errores y precisión.

Puesto que los modelos son simplificaciones de los aspectos que interesan del mundo real, la precisión de un modelo se refiere al grado de similitud existente entre el dominio fuente y el dominio final. Pero también afecta a la *resolución*, es decir, a las mediciones más pequeñas que se pueden efectuar sobre el dominio final.
La falta de precisión puede aparecer por falta de conexión entre ambos dominios, bien porque el final no sea el adecuado o porque la función de modelado no sea ajustada. Y en ambos casos tanto los resultados del análisis como las conclusiones, serían erróneos.

Por tanto, las discrepancias entre ambos dominios pueden ser de dos órdenes:

- Cualitativas: que surgen, por ejemplo, cuando existen diferencias de clasificación entre ambos dominios.
- Cuantitativas: que aparecen, por ejemplo, por diferencias de resolución, y que resultan especialmente graves en los datos espaciales porque se propagan rápidamente y de modo imprevisible.

En general, se asume implícitamente que las fases de modelado que se han comentado más arriba están totalmente libres de error: es decir, que no sólo no han existido defectos en la supervisión o la integración de los datos, sino tampoco errores estadísticos -variación mínima-, y en definitiva, que la calidad de los datos ha sido uniforme para toda el área de estudio, y que el muestreo se puede considerar completo.

Los componentes del error se pueden categorizar en tres grandes bloques: I, los que resultan evidentes y fácilmente detectables; II, los que proceden de variaciones en las mediciones originales, cuya detección requiere un conocimiento profundo de los fenómenos que se estudian; y III, los que surgen durante el proceso.
A continuación se exponen los bloques I y II, dejando el tercer grupo para posteriores estudios específicos sobre los *SIG vectoriales* y los *SIG raster*.

I. <u>Errores evidentes y fáciles de detectar</u>

- *La edad de los datos:* en general, y salvo que los datos se reúnan específicamente para un proyecto concreto, lo habitual es que procedan de diferentes fuentes y se hayan recopilado en diferentes momentos; en general se puede afirmar que, con excepción de los datos geológicos, la fiabilidad de los datos disminuye con los años. Además hay que tener en cuenta que aunque las condiciones del fenómeno que se estudia no hayan variado en el transcurso del tiempo, pueden haber cambiado los estándares de medida o los métodos, y por esta razón puede que los datos resulten también inaceptables.

- *La cobertura de los datos:* es deseable disponer de una cobertura de la información uniforme para toda la zona de estudio, pero lo cierto es que en escalas medias y grandes resulta habitual tener que trabajar con informaciones parciales, y que en estos casos hay que decidir previamente cómo conseguir la uniformidad necesaria: bien recogiendo más datos, o bien generalizando los disponibles. Resulta desaconsejable rellenar los vacíos con información procedente de mapas de escala menor.

- *La escala del mapa:* una fuente de información geográfica importante son los mapas temáticos en soporte papel. En general, las escalas grandes aportan mucha más información topográfica y leyendas más detalladas, pero lo fundamental es que la escala de los mapas-fuente coincida con los requisitos de información del SIG.

- *La densidad de las observaciones:* aún es frecuente que se editen mapas en los que no se mencionen ni los puntos en los que se han realizado los muestreos, ni las condiciones bajo las que éstos se han hecho; en consecuencia, la fiabilidad de la información que aportan es cuestionable. Aunque la densidad de observaciones puede ser un indicador sobre el grado de fiabilidad de los datos, tampoco resulta un valor absoluto, por lo que se uilizan técnicas de muestreo que utilizan el análisis de la varianza para determinar las distancias en las que aparecen los mayores aumentos de ésta. Estos estudios permiten estimar el número de muestreos necesario para definir

y trazar una zona, o para determinar los valores medios que proporcionan niveles de confianza suficientes en determinadas zonas. También es frecuente utilizar técnicas de interpolación. En general puede afirmarse que las técnicas de análisis espacial geoestadístico permiten conocer si existe información suficiente para definir las relaciones espaciales.

- *La relevancia de los datos:* no todos los datos utilizados en los SIG resultan interesantes directamente para los propósitos para los que se utilizan, pero se usan como sustitutivos ante la imposibilidad de contar con los datos adecuados. Un ejemplo lo constituyen las señales procedentes de sensores remotos que se utilizan para definir los usos del suelo. Pero si las relaciones que existen entre los datos sustitutorios y los deseables están suficientemente definidas y contrastadas -lo que por ejemplo ocupa una gran parte de la tecnología de la Teledetección-, entonces resulta aceptable tal sustitución.

- *El formato de los datos:* los errores debidos al formato de los datos surgen en los SIG de tres maneras diferentes:
    - en primer lugar, desde un punto de vista técnico aparecen relacionados con el modo como se graban los datos en un medio determinado para permitir su transferencia entre sistemas;
    - en segundo lugar, hay que tener en cuenta la manera de estructurar los datos;
    - y en tercer lugar intervienen aspectos propios de los datos tales como la escala o la proyección con las que están vinculados -cuyas conversiones se realizan fácilmente mediante transformaciones geométricas-, y la clasificación de la que resultan -que suele ser difícil de comprobar.

  Hay que reconocer que aunque existen formatos estándar para los datos procedentes de la Teledetección y para los datos topográficos entre los países de la Unión Europea, aún no existe internacionalmente una uniformidad total al respecto.

- *La accesibilidad y el coste:* no existen las mismas posibilidades de acceso a los datos en unos organismos o países, que en otros. En España se recomienda

contactar con las empresas y con los organismos cartográficos de las Administraciones Públicas para conocer el grado de accesibilidad a los datos disponibles, su formato y su coste.

También existen agencias que informan sobre los servicios de información internacionales, como la *Euronet DIANE News* de la Comisión de las Comunidades Europeas, que es accesible a través de Internet[2].

II. Errores por variación de las mediciones originales

- *Errores de posición*: que se producen en las localizaciones de los elementos espaciales del modelo, para un cierto grado de precisión, y respecto a las posiciones del dominio fuente. Su importancia depende del tipo de datos que se consideren: por ejemplo, los datos topográficos requieren un alto grado de precisión para definir los objetos geográficos -carreteras, presas, construcciones, lindes, etc.-, que hoy es fácilmente alcanzable a niveles submétricos utilizando medios como el G.P.S. –siempre que se den las condiciones adecuadas. Por contraste, la posición de los límites existentes entre unidades de cultivos o coberturas vegetales suele quedar a criterio del usuario; y además frecuentemente se interpenetran, lo que dificulta la decisión. Estos errores pueden resultar de un trabajo de campo insuficiente, de deformaciones del soporte de los mapas de trabajo, e incluso de deficiencias en la vectorización. Tienen solución aumentando el trabajo de campo o aplicando técnicas como la "hoja elástica" (*rubber sheeting*).

---

[2] *Computerwoche:* http://www.computerwoche.de

R. Haber (List of Publications from the Bibliography Server on Database Systems & Logic Programming): http://researchsmp2.cc.vt.edu

European Research and Technological Development:

http://www.cordis.lu/cordis/cord2000.html

También está disponible una dirección que proporciona todas las correspondientes a los organismos de investigación vinculados a la Euronet DIANE en:

EU:* http://www.bs.cs.tu-berlin.de/jensd/glossary/EU.html

- *Errores en el valor del atributo*: los atributos no espaciales de los elementos del modelo han de ser compatibles, para un cierto grado de precisión, con las propiedades del dominio origen. Hay que distinguir entre:
    - o precisión cualitativa, que afecta a los valores nominales -por ejemplo, al cambiar el uso de suelo Adocente@ por el de "investigación"-,
    - o y precisión cuantitativa, que se refiere a la desviación estimada en los valores obtenidos -por ejemplo, si en las mediciones se utiliza un iinstrumento mal calibrado.

- *Inconsistencia lógica*: que afecta a los elementos espaciales y no espaciales del modelo, que han de ser respectivamente consistentes; por ejemplo, manteniendo las reglas topológicas que entre los primeros existan.

- *Modelo incompleto*: el modelo debe manejar todos los elementos del dominio fuente que se necesiten.

En general no hay que olvidar que el modelo no es el mundo real, y que una semejanza muy estrecha con éste puede limitar su comprensión.
Por ejemplo, muchos modelos manejan entidades asimiladas a formas geométricas, y las formas geométricas que existen sobre la superficie terrestre son prácticamente todos de naturaleza artificial, es decir, construidas por el hombre -carreteras, lindes, etc.-, pues las formas naturales son mucho más arbitrarias y no son fáciles de definir matemáticamente.

En definitiva, los SIG construyen modelos de información espacial sobre la base de su expresión matemática y de la posibilidad de que sea procesable informáticamente. Ambos aspectos son los que se desarrollan en los apartados siguientes.

2. Los datos.

A diferencia de otros tipos de datos que manejan habitualmente los sistemas de información generales, los datos geográficos son complejos porque incorporan información sobre posición, posibles relaciones topológicas, y atributos de los fenómenos que se estudian. Por tanto, se puede afirmar que los aspectos topológicos y espaciales del proceso de datos geográficos, resultan distintivos frente a otros sistemas de información como puedan ser los utilizados en bancos o bibliotecas.

Pero para poder manipular coherentemente los datos, los fenómenos geográficos han de definirse de acuerdo con un concepto de espacio determinado, y han de referirse a posiciones sobre la superficie terrestre (geo-referenciación) utilizando sistemas de coordenadas convencionalmente aceptados, que pueden ser de ámbito local -como cuando se estudian zonas de pequeña extensión- o proyecciones internacionalmente utilizadas como la UTM (*Universal Transverse Mercator*), con sus correspondientes sistemas de coordenadas.

En definitiva, es imprescindible definir matemáticamente el espacio en el que se van a estudiar los fenómenos, de modo que tal espacio pueda modelarse y ser procesable por el ordenador.

2.1 Las concepciones de espacio y los modelos de SIG. Datos espaciales y atributivos.

Si utilizamos una definición muy amplia de espacio al considerarlo como <<una relación establecida entre un conjunto de objetos>> (Gatrell, 1991), dentro de ella tiene cabida cualquier grupo estructurado de objetos o fenómenos, e incluso la idea intuitiva del espacio físico que perciben nuestros sentidos.

En el ámbito de los SIG, es posible concebir el espacio de dos maneras diferentes:

- Como un conjunto de posiciones dotadas de propiedades -el espacio absoluto, existente en sí mismo.

- O como un conjunto de objetos con propiedades espaciales -espacio relativo, dependiente de otros objetos.

Esta dicotomía presenta claras vinculaciones con las concepciones filosóficas clásicas, y tiene consecuencias evidentes en el modelado espacial, pues la idea de espacio absoluto se modela como un conjunto de campos -es decir, de atributos-, mientras que la idea de espacio relativo se modela como conjuntos de objetos espacialmente referidos.

Ambas concepciones resultan claramente aplicables al espacio geográfico, ya que éste se puede considerar:

- Como una colección de posiciones sobre la superficie terrestre, sobre las que tienen lugar los fenómenos geográficos.
- O como una colección de entidades geográficas referidas a una parte de la Tierra, que forman parte de los procesos geográficos.

Además, cuando se habla del diseño de un SIG, las filosofías absoluta y relativa se pueden asociar respectivamente con las tecnologías SIG raster y vectorial.

ESPACIO	MODELADO	TECNOLOGÍA
Absoluto	Conjunto de Campos	Raster
Relativo	Colecciones de objetos espacialmente referidos	Vectorial

Por tanto, el *modelo de SIG basado en campos* trata los datos como colecciones de distribuciones espaciales, cada una de las cuales puede formalizarse por medio de una función matemática en una trama espacial. Por ejemplo, una malla rectangular superpuesta sobre un modelo ideal de dominio atributivo, como puedan ser las altitudes topográficas, la pluviosidad o la temperatura. Es decir, que cada campo define la variación espacial de un atributo, como una función que relaciona el conjunto de las posiciones (*trama espacial*) con un dominio atributivo. El campo es la función, no el conjunto de valores.

En cambio, el *modelo basado en objetos* trata la información espacial como si estuviese constituida por entidades discretas y definidas, cada una de las cuales está georreferida. Es decir, que construye un grupo de entidades integradas en un espacio, y son las propias funciones aplicadas a las entidades las que definen el espacio. Por tanto, en lugar de concebir la trama espacial como un "tapiz" subyacente, y de dejar que los valores atributivos varíen respecto a él -por ejemplo, las densidades de población-, la aproximación basada en objetos llena el espacio de la información con objetos -edificios, parcelas, lindes, etc.- que poseen atributos -superficie, centroide, límite, etc.- asimilables a las características de otros objetos -número racional a la superficie, punto al centroide, polígono al límite.

Ambos modelos se materializan en dos tipos diferentes y opuestos de SIG, cuales son el tipo raster y el tipo vectorial.

2.2 *La naturaleza singular de los datos espaciales. Espacios y Geometrías.*

Las geometrías proporcionan propiedades abstractas de objetos y estructuras en el espacio.
Desde que Klein enunciase en 1872 el programa *Erlangen* basado en la noción de *invariancia*, las geometrías se distinguen por ser el grupo de las transformaciones en el espacio bajo las cuales sus proposiciones permanecen verdaderas.
Por ejemplo, si consideramos un espacio tridimensional y lo que entendemos habitualmente por distancia entre dos puntos, el conjunto de transformaciones que mantienen la distancia -es decir, que para ellas la distancia es un *invariante*-, forman una geometría. En este conjunto se encuentran las traslaciones y las rotaciones, pero no el cambio de escala.

Por tanto, definimos una *geometría* como el estudio de los invariantes de un conjunto de transformaciones.

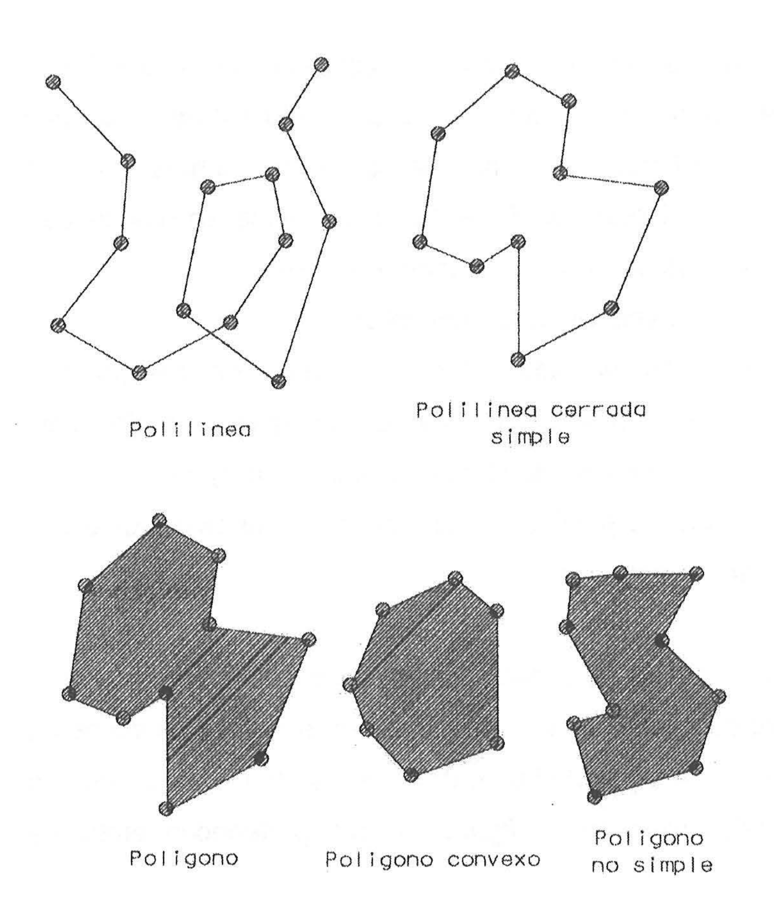
Son varias las geometrías que se utilizan habitualmente en la definición del espacio geográfico en un SIG, y entre ellas se encuentra el espacio Euclídeo, el espacio basado en la Teoría de Conjuntos, el espacio Topológico, los espacios en Red, y los espacios Métricos.
Cada uno de ellos incide en ciertos aspectos del dominio de aplicación, que facilitan el diseño del correspondiente modelo de información espacial.

2.3 *Geometrías más usuales en la tecnología SIG.*

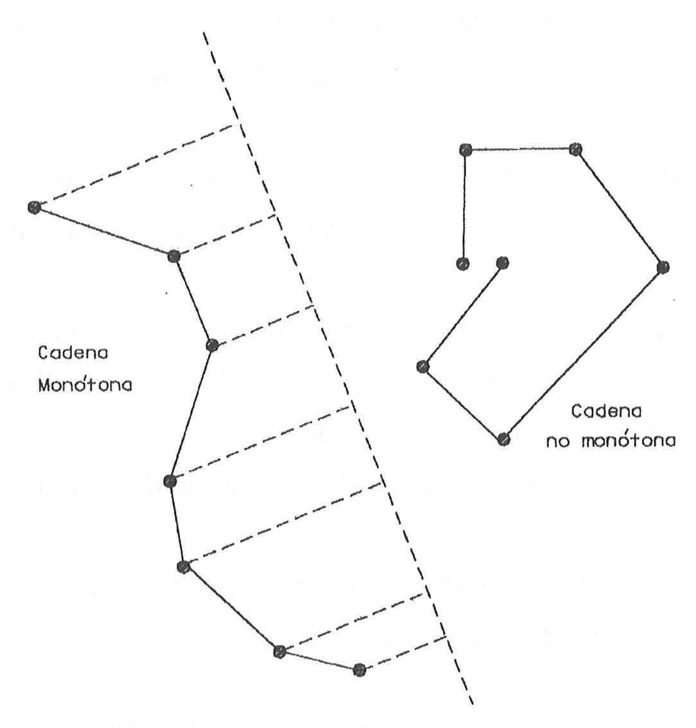
A/ El espacio Euclídeo.

Muchos fenómenos geográficos se modelan considerándolos integrados en un espacio coordenado, que permite medir distancias entre puntos utilizando ciertas expresiones matemáticas.
Un ejemplo es el espacio Euclídeo, que es un espacio coordenado que transforma las propiedades espaciales en propiedades de los números reales.
Para simplificar vamos a considerar un modelo bidimensional -el *plano euclídeo*-, con un sistema de coordenadas consistente en un punto *origen* fijo y en un par de *ejes* ortogonales que intersecan en el origen; es posible definir:

- *Objetos puntuales*: un punto en el plano euclídeo tiene asociado un único par de números reales *(x,y)*, midiendo la distancia desde el origen en la dirección de cada eje. El conjunto de todos los puntos constituye el *plano cartesiano* (R^2), y resulta habitual considerar los puntos *(x,y)* como vectores medidos desde el origen hasta el punto, dotados de dirección, sentido y módulo, y representados por un segmento orientado.
  En consecuencia, pueden ser sumados, restados y multiplicados escalarmente, pero igualmente permiten realizar operaciones vectoriales.
- *Objetos lineales*: representan los atributos espaciales de los objetos y sus contornos. Matemáticamente es posible definir la recta, la semirrecta y el segmento, así como secciones cónicas y curvas mediante expresiones polinómicas.



Polilíneas y tipos de polígonos



Cadenas monótonas y no monótonas

- *Objetos poligonales*: una *polilínea plana* se define como un conjunto finito de segmentos lineales (llamados *lados*) tales que cada punto extremo de cada segmento pertenezca a dos lados, con la salvedad de los dos puntos extremos de la polilínea. Si además ningún par de lados se interseca en ningún otro punto que en sus extremos, la polilínea se denomina *simple*.

  Una polilínea es *cerrada* si carece de puntos extremos.

  Un *polígono simple* en R^2 está definido por la superficie encerrada por una polilínea simple cerrada; ésta forma el *contorno* del polígono, y cada punto extremo de los lados de la polilínea se denomina *vértice* del polígono.

  Generalizando este concepto, es posible que un polígono contenga "agujeros", e incluso "islas" dentro de los agujeros.

Los programas de SIG, como los de CAD, permiten definir varios tipos de polígonos:

- *Convexos*, cuyo interior es un conjunto convexo, por ser sus ángulos interiores menores de 180º [3] Poseen la propiedad de que todos los puntos del segmento que une dos puntos cualesquiera de una figura convexa, pertenecen también a ella.

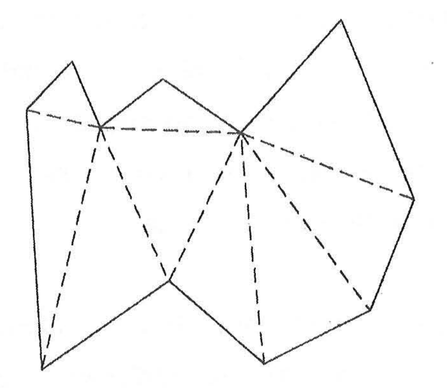
  En los SIG resultan muy útiles, por ejemplo, para analizar la visibilidad dentro de un área geográfica, ya que en virtud de esta propiedad cada punto interior es visible desde cualquier otro punto interior: la línea de visión es completamente interior al polígono.
- *No simples*, son los polígonos no convexos que poseen la propiedad de que debe existir al menos un punto que sea visible desde cualquier punto del polígono.
- *Monótonos*, son aquéllos que se basan en el concepto de *cadena monótona*: una cadena C =$|p_1, p_2, p_3, ...p_n|$ es un conjunto ordenado de *n* puntos en el plano euclídeo, y es monótona si, y sólo si, existe alguna línea en el plano tal que la proyección de los puntos sobre la línea mantenga el orden de la lista de puntos.

Un polígono es monótono si su contorno puede descomponerse en dos polilíneas tales que la cadena de vértices de cada polilínea sea monótona.

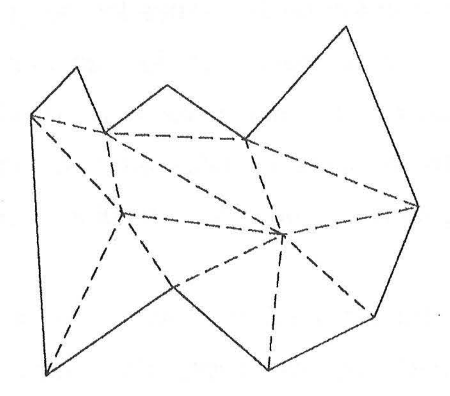
[3] Dados en un determinado orden *n* puntos del plano que no están alineados, si las rectas determinadas por cada dos puntos consecutivos dejan en un mismo semiplano los *n-2* puntos restantes, el conjunto de puntos comunes a todos estos semiplanos es un polígono convexo.

Todo polígono convexo es monótono, pero no a la inversa. Tampoco todo polígono monótono es estrellado.

Los polígonos pueden *triangularse*, es decir, subdividirse en triángulos que sólo se intersecan en los vértices de sus respectivos contornos. Una triangulación de un polígono simple de *n* vértices, que incorpora *m* vértices de polígonos internos por triangulación, proporciona *n* + *2m* -*2* triángulos. Si no se introducen vértices interiores (*puntos de Steiner*), la triangulación se denomina *diagonal*, y proporciona *n* - *2* triángulos.



Triangulación diagonal



Triangulación no diagonal

Estos conceptos -y sus algoritmos respectivos- resultan de gran utilidad en los SIG cuando se desea representar un modelo a base de mosaicos regulares o irregulares, como por ejemplo una *red irregular de triangulación* (TIN, *Triangulated Irregular Network*), que se aplica en estudios de pendientes, de impacto ambiental, en la determinación de perfiles topográficos, etc.

Por otra parte, una *transformación* en R^2 es una función que aplica el plano euclídeo en sí mismo, de modo que todo punto del plano se transforma en otro -o el mismo- punto.

Las transformaciones que mantienen ciertas propiedades particulares de los objetos integrados en el plano euclídeo, y que se utilizan con mayor frecuencia en un SIG son:

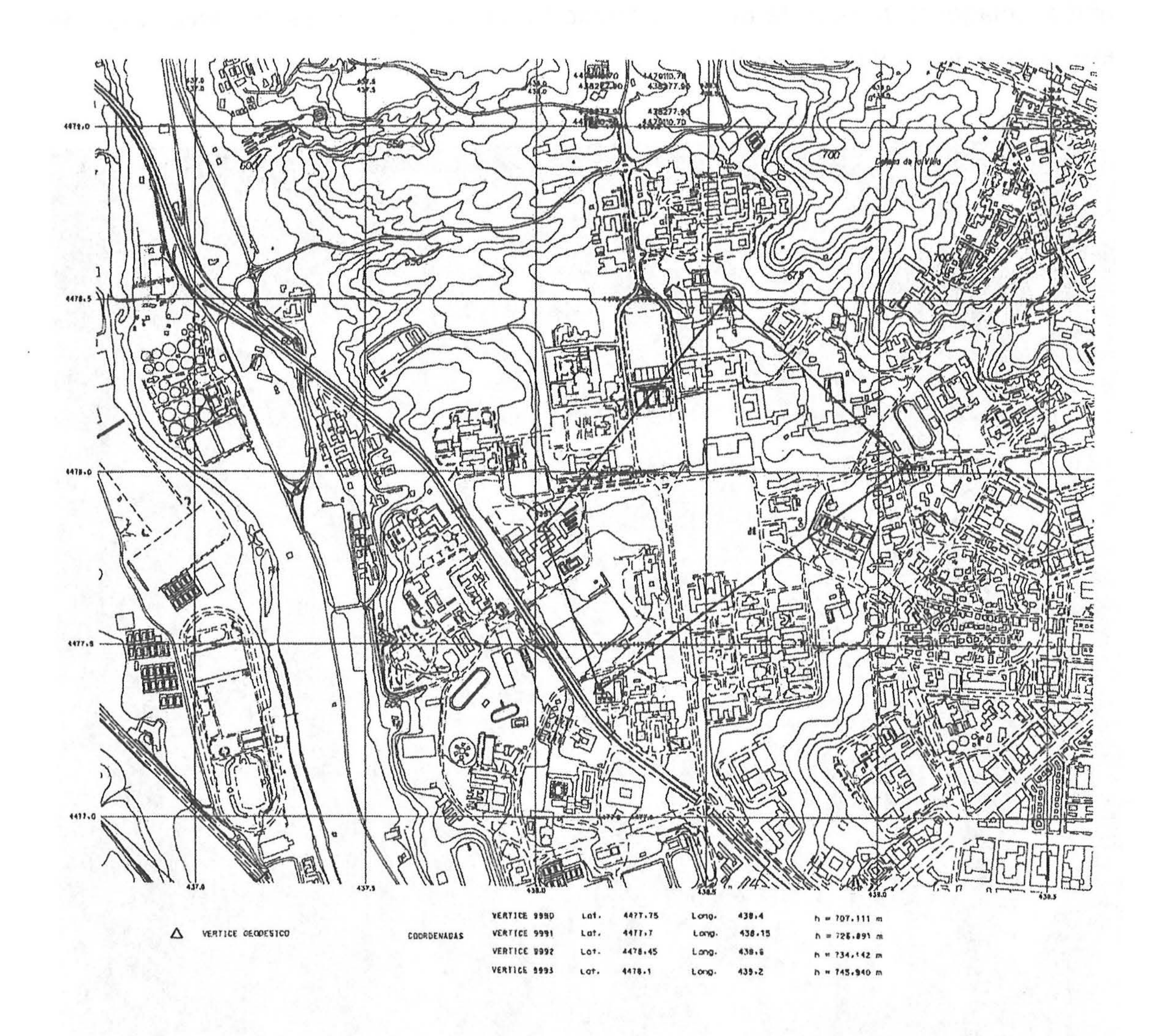
- Las *transformaciones euclídeas (congruencias)*, que permiten obtener una figura de otra aplicándole a ésta un movimiento; mantienen la forma y el tamaño de los objetos -por ejemplo, la traslación.

- Las *transformaciones de semejanza*, que aplicadas a una figura obtienen otra, de modo que entre los puntos de ambas se pueda establecer una correspondencia biunívoca que cumpla que: 1º a puntos alineados corresponden puntos alineados de igual orden; 2º los segmentos homólogos son proporcionales; 3º los ángulos homólogos son iguales. Por tanto, mantienen la forma, pero no necesariamente el tamaño, como sucede en el cambio de escala. Todas las transformaciones euclídeas son semejantes -la congruencia es un caso particular de la semejanza, de razón unidad.

- Las *transformaciones afines (homologías afines):* dos formas planas perspectivas, es decir, dos secciones de una misma radiación, son homológicas; si además el eje de la homología es propio, y el centro impropio, la homología es afín. Por ello, una homología afín queda definida dando el eje y un par de puntos homólogos. Mantienen propiedades como la razón simple, el paralelismo, el giro, la simetría o el desplazamiento.

- Las *transformaciones proyectivas:* dos formas planas son proyectivas entre sí cuando pueden obtenerse una de otra por una sucesión de proyecciones y secciones; mantienen propiedades tales como la razón doble, la polaridad, o la ordenación -aunque según el tipo de proyectividad existen muchas más, como la incidencia, la intersección y tangencia, etc. Todas las transformaciones afines son también proyectivas.

- Las *transformaciones topológicas (homeomorfismos)*, que apoyándose en la teoría matemática de la descomposición celular de las superficies continuas, mantienen las propiedades de las posiciones relativas de los elementos -conjunto abierto o cerrado, contorno, etc.

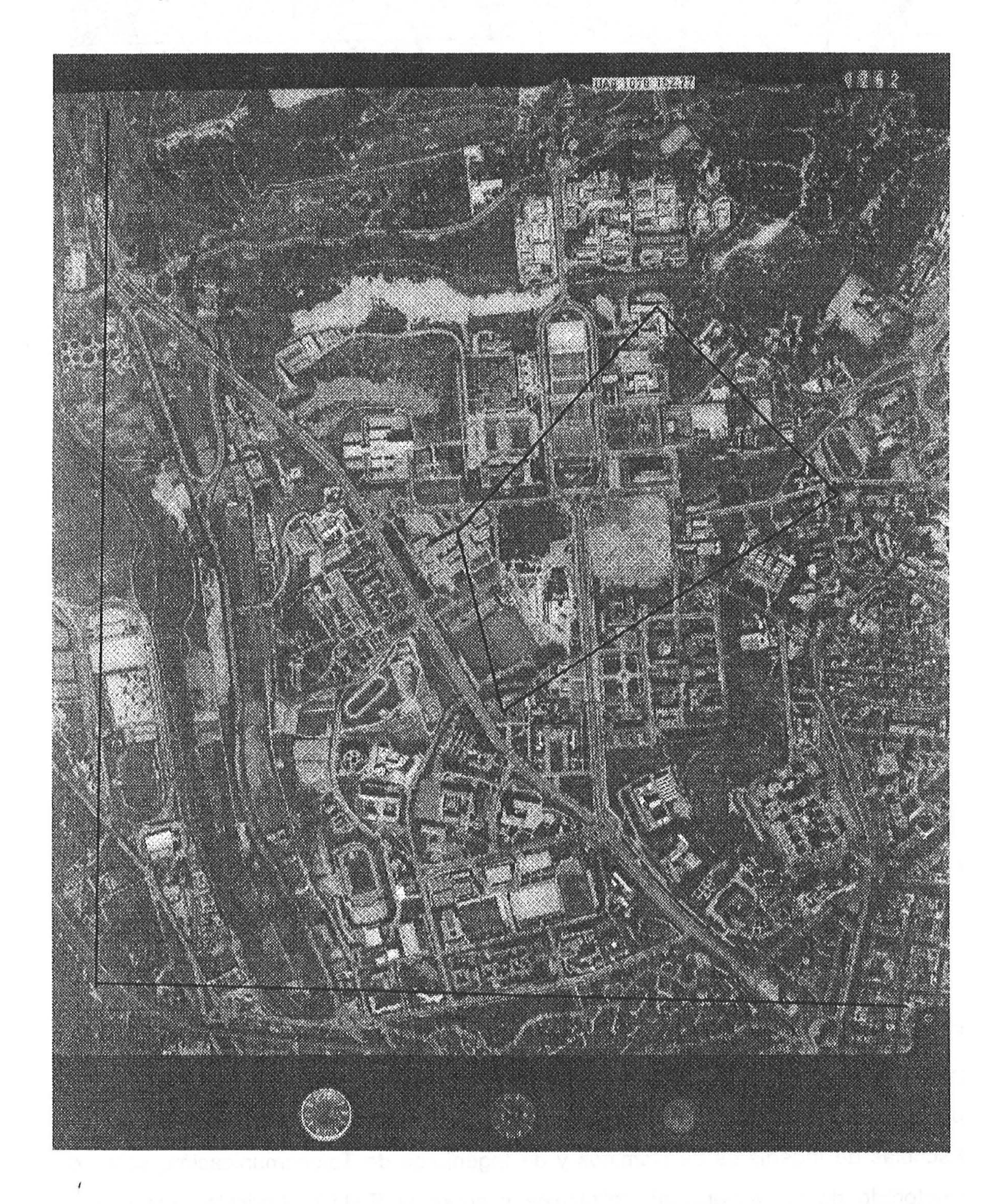


La utilidad que para los SIG tienen estas transformaciones también resulta inmediata. Por ejemplo, supongamos que disponemos de dos documentos gráficos:

- una fotografía aérea vertical de la Ciudad Universitaria no orientada, de la que sólo se conoce la escala aproximada;
- y un plano a escala 1:5.000 que representa el mismo ámbito geográfico.

Para trasladar el mismo sistema coordenado del plano a la fotografía será necesario partir de una serie de puntos de control de coordenadas conocidas en el sistema del plano -por ejemplo, los vértices geodésicos situados sobre las cubiertas de las Escuelas de Ingenieros de Caminos y de Ingenieros de Telecomunicación, sobre el Rectorado de la Universidad Politécnica y sobre la E.U.I.T. Agrícola, que unidos forman el polígono que aparece en la ilustración-, y después efectuar las correspondientes transformaciones de giro de los ejes coordenados -en este caso

ambas imágenes poseen la misma orientación, por lo que no resulta necesario-, de traslación y de cambio de escala.

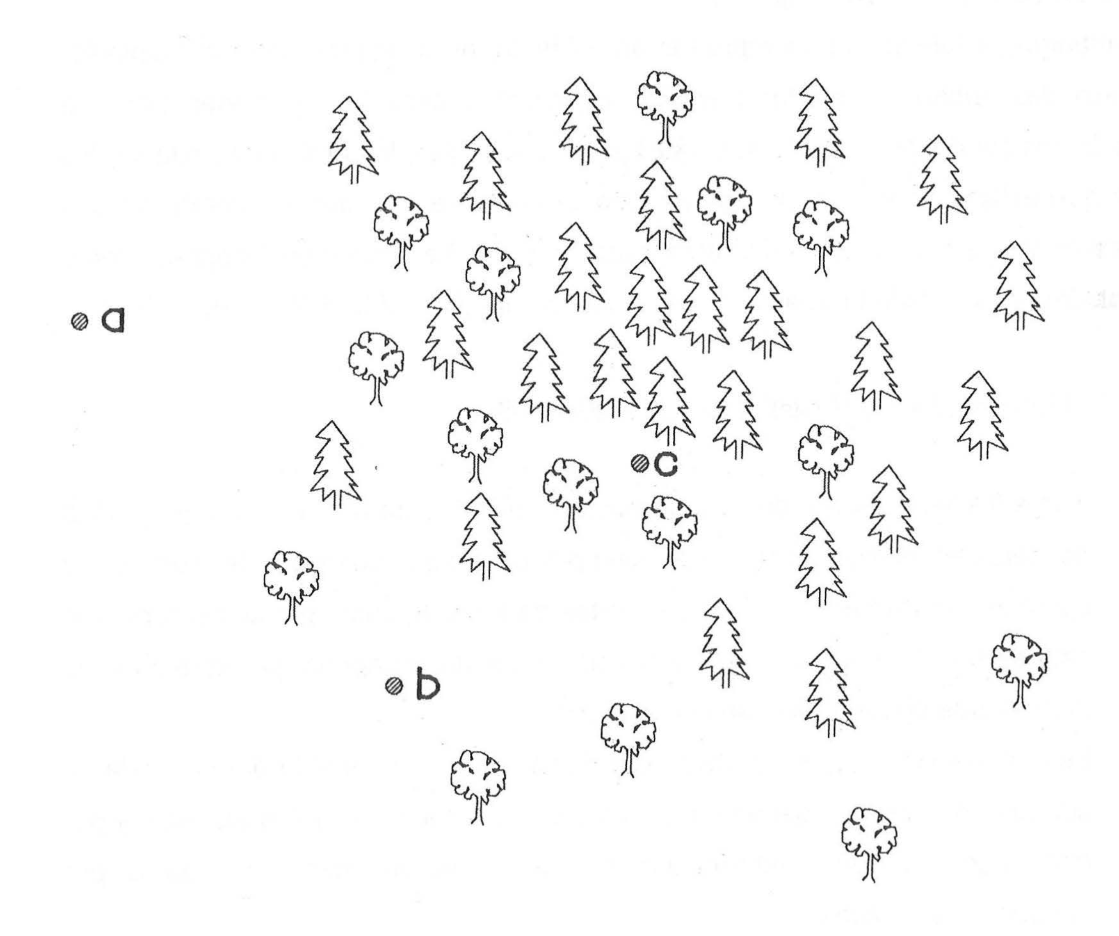


Estas transformaciones permitirán conocer la posición exacta de cualquier punto de la fotografía.

B/ La Geometría del Espacio basado en la Teoría de Conjuntos

Un modelo de espacio basado en la Teoría de Conjuntos no posee la riqueza de construcciones que proporciona un espacio euclídeo, ya que básicamente utiliza *elementos, conjuntos* de elementos y la relación de *pertenencia* entre un elemento y el conjunto al que se aplica. Expresa, por tanto, las relaciones más primitivas y la estructura más primaria de espacio.

En los SIG, son *elementos* o *miembros* los objetos constituyentes que se van a modelar, que tienen que estar bien definidos y ser discernibles. Son *conjuntos* las colecciones de elementos a modelar; y para que puedan ser manejados por los ordenadores, son generalmente finitos (o al menos se tienen que poder contar). Finalmente, la relación de *pertenencia* es la que se establece entre los elementos y los conjuntos en los que se integran.



En la teoría clásica, un objeto o es un elemento, o es un determinado conjunto, o no lo es; no existen términos medios, ni grados de pertenencia. Pero actualmente se tiende a relajar, por razones prácticas, la naturaleza binaria de esta relación de pertenencia, y se plantea el aplicar el modelo de conjuntos de la *lógica fuzzy* o *difusa* "inventada" en 1988 por Zodeh. Este modelo permite resolver situaciones que sepresentan habitualmente en los SIG, como por ejemplo la proposición: "Posición *x* en área de bosque": si consideramos sólo 1/verdadero y 0/falso, pueden darse situaciones en las que resulte difícil optar por uno de los dos valores porque podrían asignarse valores verdaderos entre el 0 y el 1. Tal es el caso de la posición *b* de la figura..

En general, este modelo de espacio es muy abstracto y aporta pocas estructuras para definir propiedades y relaciones espaciales, pero resulta esencial en el modelado de cualquier sistema de información espacial. Por ejemplo, la jerarquía que existe entre los conceptos de *ciudad - provincia - comunidad*, se define o modela adecuadamente por medio de la Teoría de Conjuntos.
Sin embargo, la relación no siempre es tan evidente, ni resulta tan sencillo manipular las unidades superficiales. Por ejemplo, los distritos censales no tienen por qué coincidir con los distritos postales, y este hecho causa grandes problemas cuando los datos que están referidos a un determinado conjunto se comparan o combinan con datos referidos a otro conjunto diferente. Esto es lo que se conoce en Geografía como el Problema de la Unidad Espacial Modificable (*Modified Areal Unit Problem, MAUP*).

El Problema de la Unidad Espacial Modificable.

Surge frecuentemente cuando se utilizan unidades de observación geográfica de carácter artificial, como por ejemplo diferentes unidades territoriales, y aparece precisamente porque los límites de los elementos geográficos no son naturales ni fíjos -como pudiera ser una divisoria de aguas, por ejemplo-, de manera que pueden variarse con facilidad.
Esta modificación, que no altera la existencia de una variable determinada, sí que puede repercutir seriamente en los valores que tal variable alcanza, y por tanto puede producir información errónea si se manipulan los datos por métodos estadísticos.

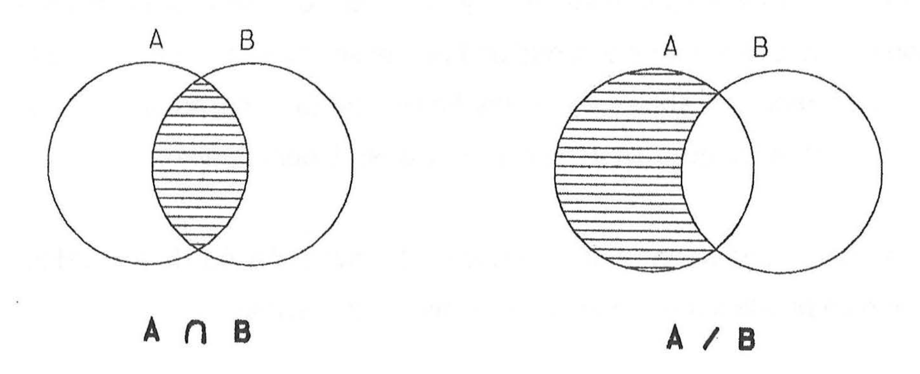
Un ejemplo muy claro es el de la alteración del ámbito de aplicación superficial de la variable "uso": el campo o variable temática sigue existiendo tanto en el estado inicial como en el modificado, pero cualquier cálculo o análisis realizado sobre los nuevos datos difiere de los primitivos y puede llevar a conclusiones erróneas. Por ejemplo, el variar los límtes superficiales del ámbito de estudio puede llevar a la conclusión errónea de que las superficies destinadas en el planeamiento a dos usos diferentes -por ejemplo, el docente o el de parque urbano- son equivalentes, cuando la realidad puede ser muy otra. Imagínese la repercusión que tendría trabajar sobre los límites de la finca primitiva de la Universitaria, o sobre los que le asigna el Plan General, por ejemplo.

Por tanto, no sólo puede constituir un error en la manipulación de los datos, sino que también puede ser un recurso para falsear la realidad.

Volviendo a las propiedades que aporta la Teoría de Conjuntos a un modelo espacial de datos, podemos resumirlas en:

- *Igualdad*: relación que se produce entre dos conjuntos cuando poseen exactamente los mismos elementos ($A = B$).
- *Subconjunto*: es la relación que existe entre dos conjuntos, cuando todo miembro de uno de ellos (A) es miembro del segundo (B); se dice que A está contenido en B ($A \subseteq B$).
- *Conjunto vacío*: es el que carece de elementos (Φ).
- *Complementariedad*: es la operación que aplicada a un conjunto A de un conjunto de referencia U, devuelve el conjunto de elementos de U que no son del conjunto (A').
- *Producto*: es la operación que a partir de dos conjuntos produce el conjunto de pares ordenados cuyos primeros elementos son miembros del primer conjunto, y cuyos segundos elementos son miembros del segundo conjunto ($A \times B$).
- *Intersección*: es la operación que se aplica sobre dos conjuntos dados, y que proporciona el conjunto de elementos que son miembros de ambos ($A \cap B$).

- *Unión*: es la operación que de dos conjuntos dados produce un conjunto cuyos elementos son miembros de, al menos, uno de los dos conjuntos iniciales (A U B).
- *Diferencia*: es la operación que, aplicada a dos conjuntos, devuelve el conjunto de los elementos que pertenecen al primer conjunto, pero no al segundo (A / B).



Una operación muy utilizada en los SIG es el *producto*, cuyo ejemplo más claro lo constituye el conjunto de puntos del plano cartesiano: cada punto se representa como un par ordenado de números reales, midiendo su distancia desde un origen dado en la dirección de los ejes; y la colección de puntos es un conjunto producto del conjunto de los números reales consigo mismo (R x R, R^2). Es un concepto generalizable a los espacios cartesianos tridimensionales (R^3) o multidimensionales (R^n).

Por ejemplo, el conjunto de las coordenadas *(x,y)*, en las que *x* e *y* son números reales, es uno de los sistemas de referencia más utilizados en los SIG (*sistema de coordenadas rectangulares planas*). Obviamente el concepto es extensible a R^3, y en los SIG cuyo ámbito sea terrestre (*sistemas globales de información*) puede resultar útil, por ejemplo, definir un sistema cuyo origen sea el centro de la Tierra y el plano *XY* sea el ecuatorial.

Los espacios producto proporcionan a su vez nuevas definiciones de relaciones entre objetos, cuales son las *binarias*. Una relación binaria en un conjunto A es el

subconjunto G del producto cartesiano A x A, cuyos pares ordenados muestran las relaciones que existen entre los miembros del conjunto. G es el *grafo* de la relación.



Por ejemplo, si volviendo a la Universitaria consideramos el conjunto de los centros docentes de cada una de las tres diferentes universidades, podemos establecer las relaciones siguientes:

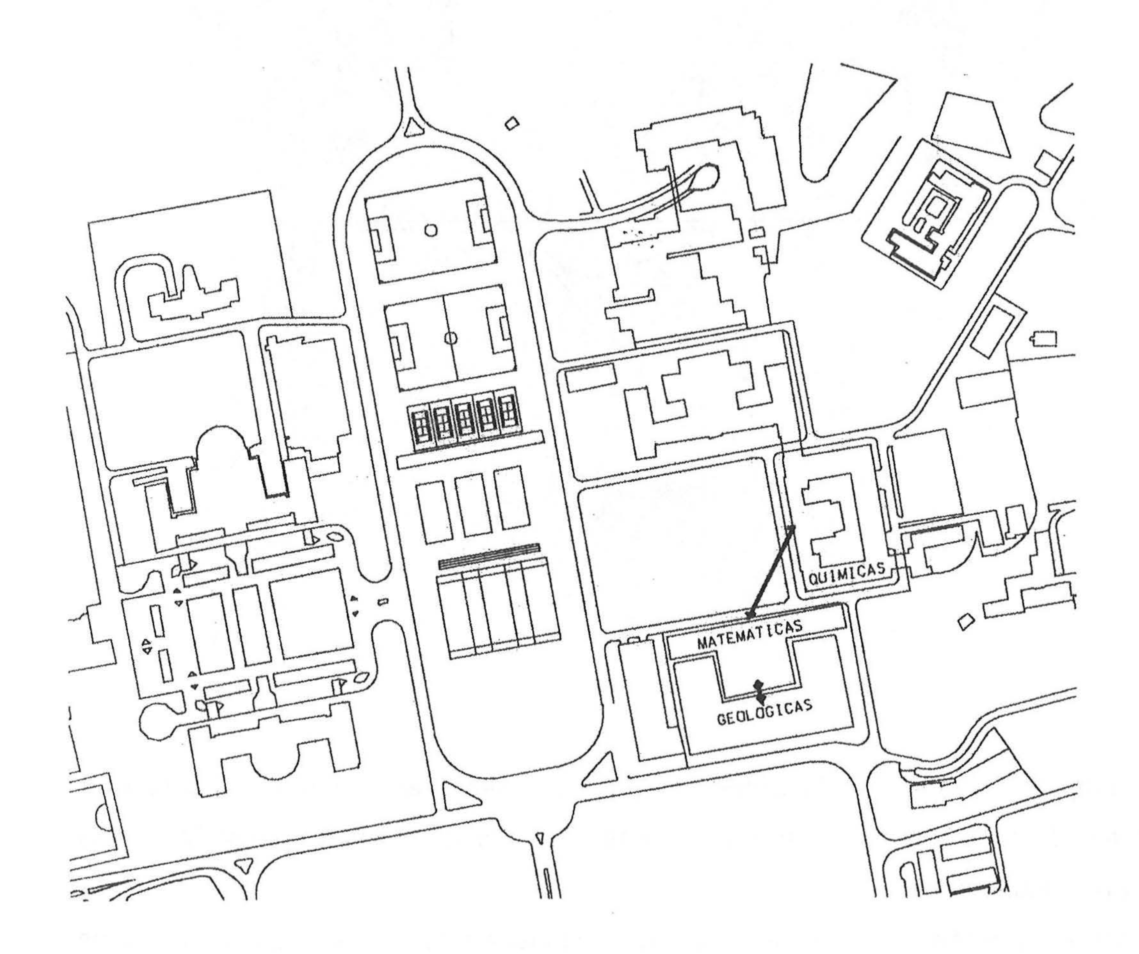
- Primera relación: supongamos que todo centro docente *d* se relaciona con un rectorado *r*, si *d* depende de *r (*correspondencia o función entre el conjunto de los centros D y el de los rectorados R).

- Segunda relación (sólo sobre el conjunto de los centros docentes): un centro d_1 se relaciona con otro d_2 , si d_2 es el centro más próximo a d_1 (relación binaria en D).

Existen además ciertas relaciones binarias que poseen ciertas propiedades especiales:

- La *relación reflexiva*, cuando cada elemento del conjunto se relaciona consigo mismo.
- La *relación simétrica*, cuando si *x* se relaciona con *y*, *y* se relaciona con *x*.
- La *relación transitiva*, cuando si *x* se relaciona con *y*, e *y* se relaciona con *z*, entonces *x* se relaciona con *z*.

Cuando se verifican las tres, la relación binaria es *equivalente*.



Siguiendo con el ejemplo de los centros docentes, y aparte de lo trivial que suponga el que un centro se relacione consigo mismo -lo que equivaldría a constatar que cada centro es el más próximo a sí mismo-, lo que no se verifica tan claramente es su

simetría: por ejemplo, supongamos que la Facultad de Ciencias Matemáticas es la más cercana a la Facultad de Ciencias Químicas: pero la Facultad de Ciencias Geológicas está más próxima a la Facultad de Matemáticas que la de Químicas. Tampoco es evidente la transitividad.

Otro tipo de relaciones habituales son las de *orden* que son relaciones binarias en un conjunto A que satisfacen la transitividad y la reflexividad, pero que son *antisimétricas* (es decir, que si $a \text{ R } b$ y $b \text{ R } a \Rightarrow a = b$); entonces el conjunto A es *ordenado*. Si dos elementos $a \text{ R } b$, se dice que "a es anterior a b" ($a \preceq b$).

Dos elementos *a,b* de A son comparables si están relacionados.
Si en una relación de orden cualquier pareja de elementos son comparables, existe un *orden total*. En caso contrario, el orden es *parcial*.

Son éstas relaciones que se aplican frecuentemente en los SIG para establecer ciertos niveles de jerarquización entre elementos espaciales. Son ejemplo de ello las estructuras "en árbol", que se presentan en los cuadernos siguientes.

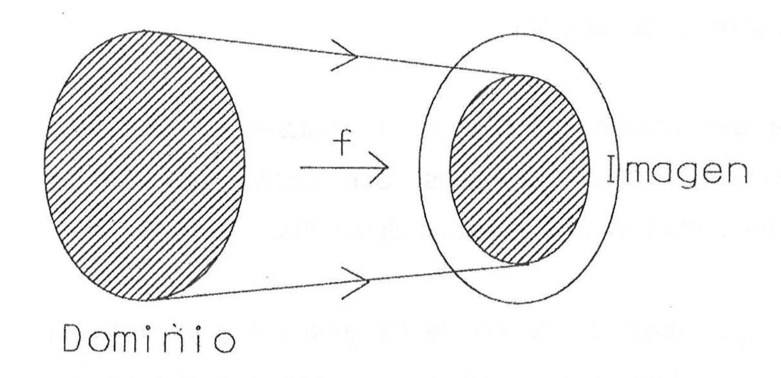
Las tres últimas propiedades que aporta la Teoría de Conjuntos son la unión, la intersección y la diferencia, y se denominan *funciones booleanas o lógicas*. Su aplicación en los SIG es frecuentísima en las operaciones lógicas que se realizan en los análisis espaciales; por ejemplo, para averiguar qué superficies o atributos verifican determinadas condiciones.
En efecto, si se quisieran conocer las superficies de la Ciudad Universitaria que mostraran la "susceptibilidad del terreno a sufrir erosión", habría que partir del dominio de los tipos de vegetación, y del de los tipos de suelo, y superponerlos.

TIPOS DE VEGETACIÓN	TIPOS DE SUELO
Arbolado denso	Arenas
Vegetación dispersa	Margas
Jardín	Aluvión
Suelos sin tratar	

La operación "Suelos sin tratar" y "Margas" proporcionará las superficies más susceptibles de erosión.

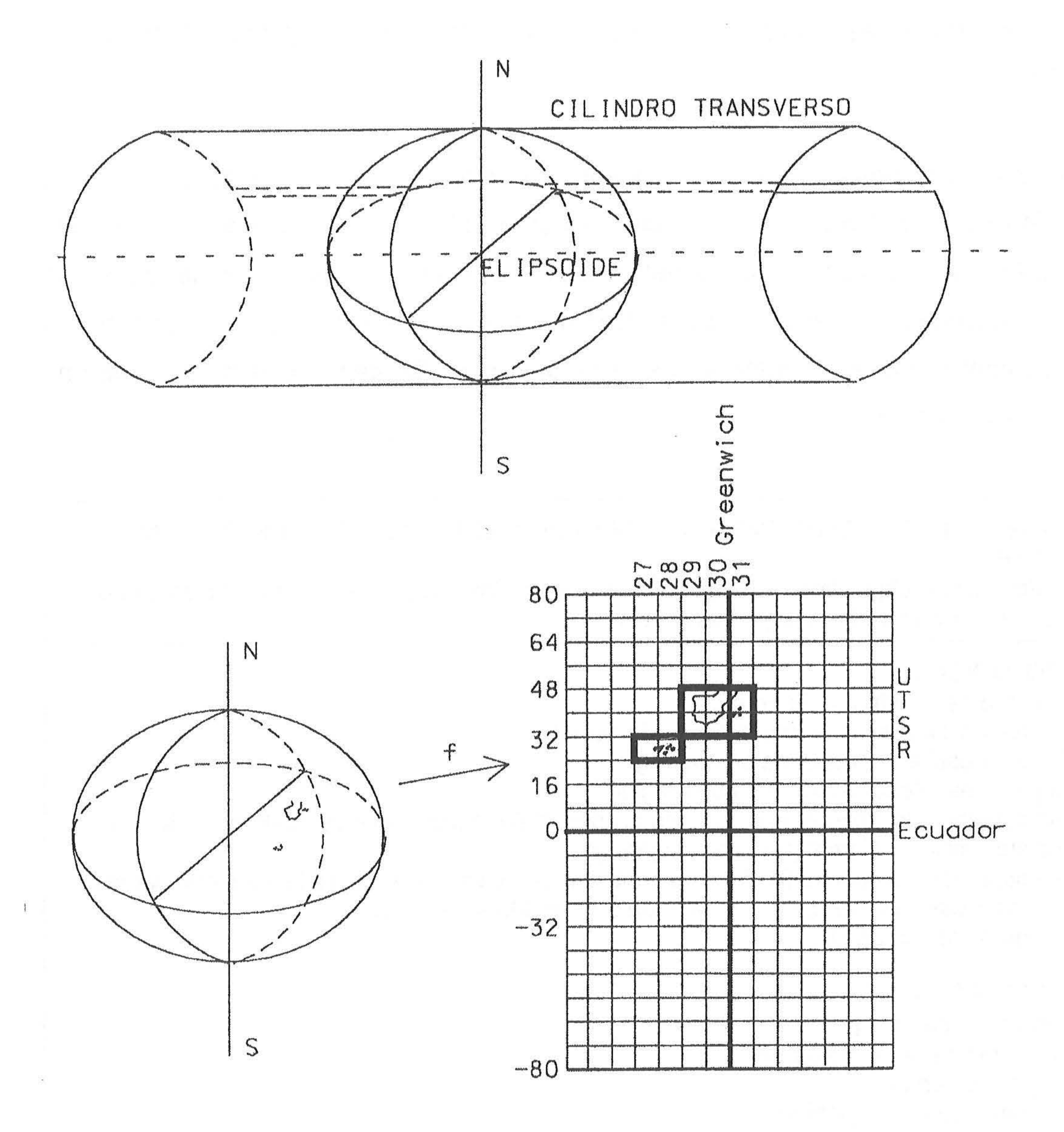
Por último, se puede definir la *función* como un tipo especial de correspondencia o transformación entre dos conjuntos A y B por la cual cada miembro del primer conjunto se relaciona exactamente con un miembro del segundo (*unívoca*). Es decir, una relación que transforme cada elemento del primer conjunto (*dominio*) en un miembro del segundo -la *imagen* de cualquier elemento de A es un conjunto unitario o vacío.

Si además, su función inversa también es unívoca, tal relación se denomina *biunívoca* o *biyectiva*.



Por ejemplo, sea A el conjunto de los puntos de un elipsoide que representa la Tierra, y B el conjunto de los puntos del plano. Las proyecciones cartográficas constituyen el grupo de funciones cuyo dominio es A y cuya imagen -conjunto resultante- es un subconjunto de B.
En general, estas funciones se construyen con el fin de mantener invariantes algunas propiedades de los conjuntos, como puedan ser las distancias, los ángulos o las áreas.

Por ejemplo, la proyección UTM (Mercator Universal Transversa) es una función que mantiene los ángulos (es *conforme*), y mediante la cual dos puntos distintos del dominio (el elipsoide) se transforman en dos puntos distintos en el plano.
Además, si aplicamos la función inversa será posible reconstruir el elipsoide a partir de su imagen cartográfica UTM.



Análogamente, aquéllas funciones en las que el conjunto imagen coincide con la totalidad del conjunto final -en el ejemplo, con todo el plano- y que además son *inyectivas* (es decir, si $f(a) = f(a') \Rightarrow a = a'$) se denominan *biyecciones*.

C/ <u>El Espacio Topológico</u>

Etimológicamente, la Topología como "estudio de la forma" constituye una parte de la Geometría que estudia el conjunto de propiedades que se mantienen invariantes bajo

las transformaciones topológicas, como por ejemplo, el ser un punto interior a una región.

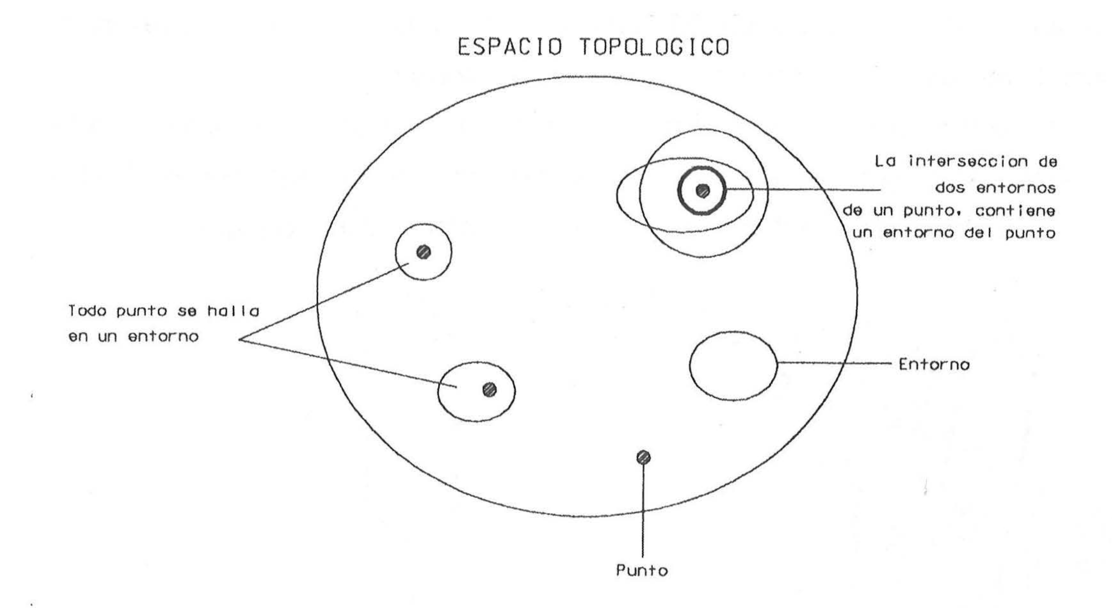
Para entender intuitivamente el concepto, se suele recurrir a imaginar el plano euclídeo como si fuera una superficie de goma sobre la que se hubieran dibujado varias formas geométricas. Si se estira el plano, algunas de las propiedades de las figuras originales se perderán, pero otras no: el punto que era interior a una de las formas continuará en su interior tras la deformación. En cambio, el área no es una propiedad topológica.

PROPIEDADES TOPOLÓGICAS Y NO TOPOLÓGICAS DE OBJETOS* EN EL PLANO EUCLÍDEO * Puntos, arcos (objeto lineal), polígonos (arcos cerrados), regiones (partes del plano, que pueden tener en su interior agujeros e islas)
TOPOLÓGICAS - Un punto es el extremo de un arco. - Un arco es simple (no se cruza sobre sí mismo) - Un punto está en el contorno de una región. - Un punto es interior o exterior a una región. - Una superficie es abierta (excluye todo el contorno) o cerrada (incluye todo el contorno). - Una superficie es simple (no tiene agujeros). - Una superficie es conexa (dados dos puntos interiores a ella, es posible seguir un camino entre uno y otro que siempre se mantenga interior a la superficie). - Un punto está en un lazo. NO TOPOLÓGICAS - Distancia entre dos puntos. - Traslación de un punto a otro. - Longitud de un arco. - Perímetro de una superficie. - Área.

En general, un espacio en el que esté bien definida la relación de *vecindad* o *entorno*, es un espacio topológico. Si A es un conjunto de puntos, un espacio topológico es una colección de subconjuntos de A, llamados *entornos*, que satisfacen las siguientes condiciones:

1º Cada punto de A está dentro de algún entorno.

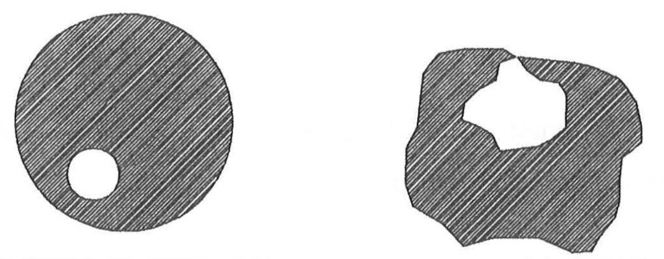
2º La intersección de dos entornos cualesquiera de cualquier punto *p* en A, contiene un entorno de *p*.



En los SIG se suele utilizar la topología usual de los puntos del plano euclídeo (que siempre puede generalizarse a un espacio euclídeo n-dimensional), y en el que un entorno de un punto es un conjunto que contiene un círculo abierto al que pertenece el punto. En efecto, una transformación topológica en R^2 transforma cada entorno del dominio en un entorno de la imagen.



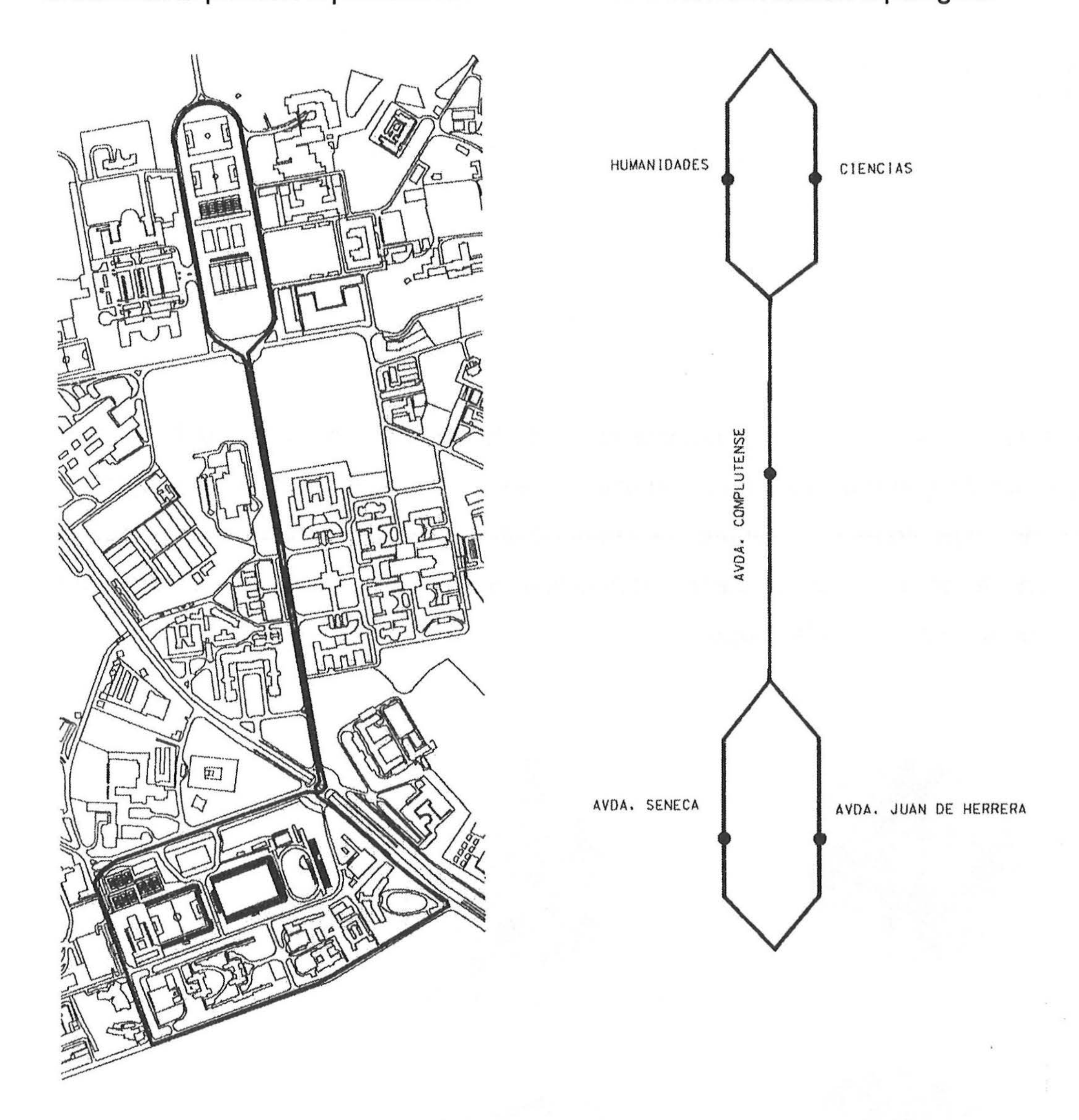
OBJETOS PLANOS TOPOLOGICAMENTE EQUIVALENTES



OBJETOS PLANOS TOPOLOGICAMENTE NO EQUIVALENTES

Y si se aplica una transformación topológica a un conjunto de puntos A y resulta el conjunto de puntos B, A y B son *topológicamente equivalentes*.

Es éste un concepto muy útil en el que se basan muchas de las representaciones cartográficas. Por ejemplo, las rutas de autobuses que circulan por la Ciudad Universitaria pueden representarse mediante una transformación topológica.



Otro de los ejemplos más frecuentes de representación topológica equivalente lo constituyen los esquemas de las redes de Metro.

Los invariantes topológicos son las propiedades que se mantienen tras efectuar una transformación topológica, como por ejemplo:

- La configuración de los entornos de un espacio.

- Los conjuntos abiertos o cerrados.
- El contorno.
- Y la conectividad: definiéndose un conjunto *conexo* como aquél en el que dos puntos cualesquiera del conjunto se pueden unir por un camino que es totalmente interior al conjunto.

Como se ha visto en el ejemplo, un arco simple es topológicamente equivalente a un segmento recto, y además es claramente conexo. Y como la transformación topológica de un segmento recto en un arco simple es una biyección, es imposible que un arco simple se cruce sobre sí mismo o que coincidan sus puntos extremos.
Si éstos coincidiesen se obtendría un lazo simple (curva cerrada simple), que es topológicamente equivalente a una circunferencia.
Precisamente sobre esta base definió Jordan su célebre *Teorema de la Curva* en 1887, que dice que toda curva cerrada plana de Jordan divide los restantes puntos del plano en dos regiones, una interior y otra exterior, tales que:

ambas están separadas por el contorno,
ambas son conexas,
y la región interior es finita pero la exterior no lo es.

Este es el fundamento de la operación "punto en polígono", que se aplica en el análisis espacial en los SIG para averiguar si un elemento puntual de un mapa se halla dentro de un área de influencia (polígono) de otro mapa. Por ejemplo, en la Universitaria podría usarse para averiguar en qué barrios vecinos -mapa de barrios- están situadas las gasolineras más próximas -mapa de gasolineras.

ALGUNOS CONJUNTOS CONEXOS EN R^2
- El plano euclídeo. - Un semiplano. - Un segmento recto. - Dos segmentos que se intersecan. - La recta. - El círculo. - La celda (objeto bidimensional primario, topológicamente equivalente a la circunferencia y simplemento conexo). - Dos celdas "pegadas" por un único punto de sus contornos. - Dos celdas conectadas por un segmento. - El anillo...

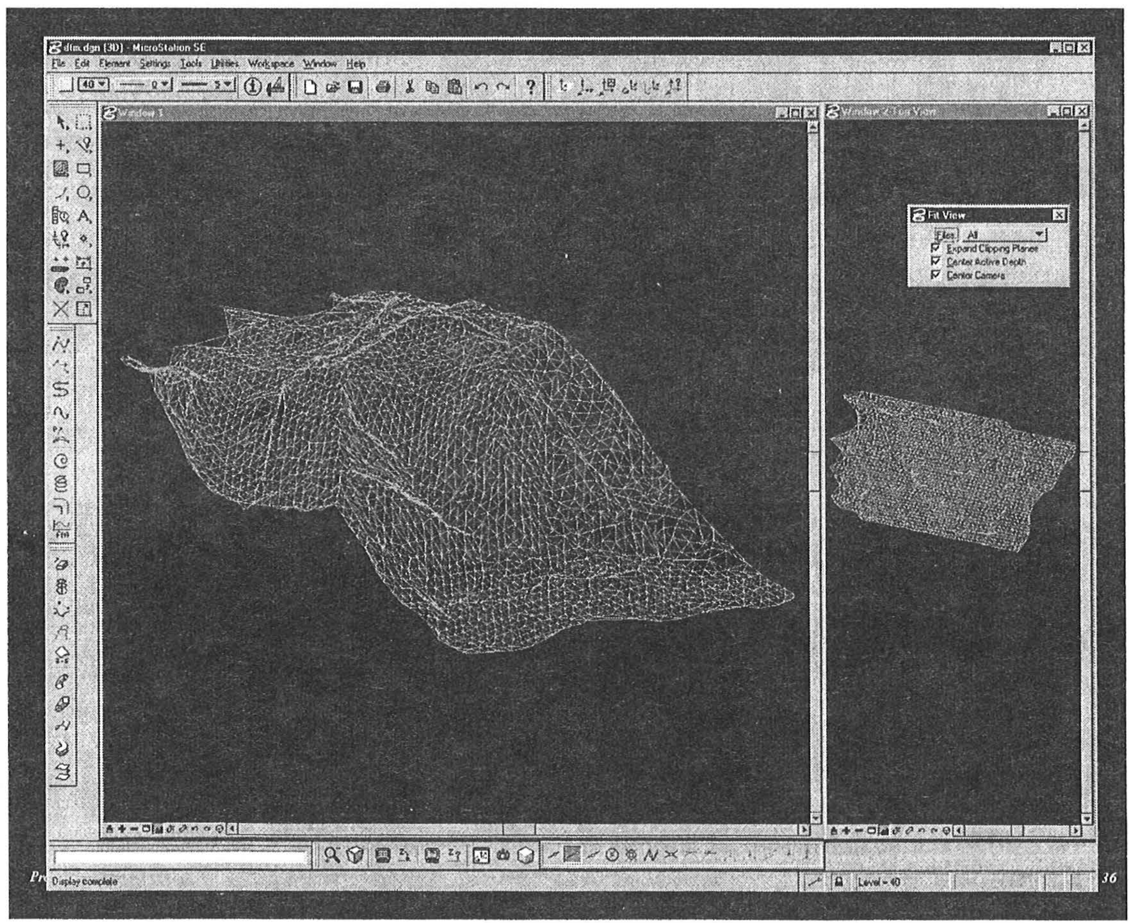
Una de las posibilidades de modelado espacial que ofrecen los SIG, cual es el orientado a objetos, se apoya en la categorización de objetos planos sobre la base de las diferentes posibilidades que ofrece la conectividad. En este caso, el manejo de celdas resulta básico para cualquier trabajo que se base en propiedades topológicas.

Si trasladamos al plano la fórmula de los poliedros de Euler $c - l + v = 2$, donde c son las caras, l los lados y v los vértices, el resultado es una configuración de celdas cuyos contornos están formados por arcos, y sus puntos de intersección son nodos; entonces se verificará que

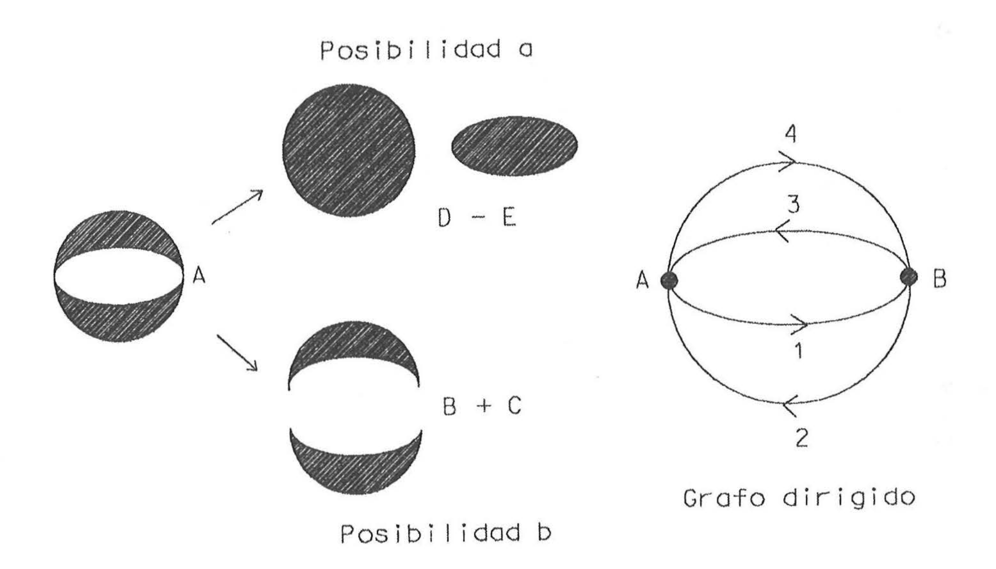
$$c - l + v = 1$$

Gran parte del trabajo con modelos genéricos del espacio se apoya en variaciones de las combinaciones de celdas planas -generalizables para las 3D.

Dos ejemplos de ello son las mallas triangulares y los grafos, ambos muy utilizados en las ciencias que trabajan con datos espaciales. En los SIG, las primeras se usan para definir los TIN (redes de triángulos irregulares), que son una manera de representar la altimetría del terreno. Los segundos son muy útiles, por ejemplo, en el estudio de redes de infraestructuras.



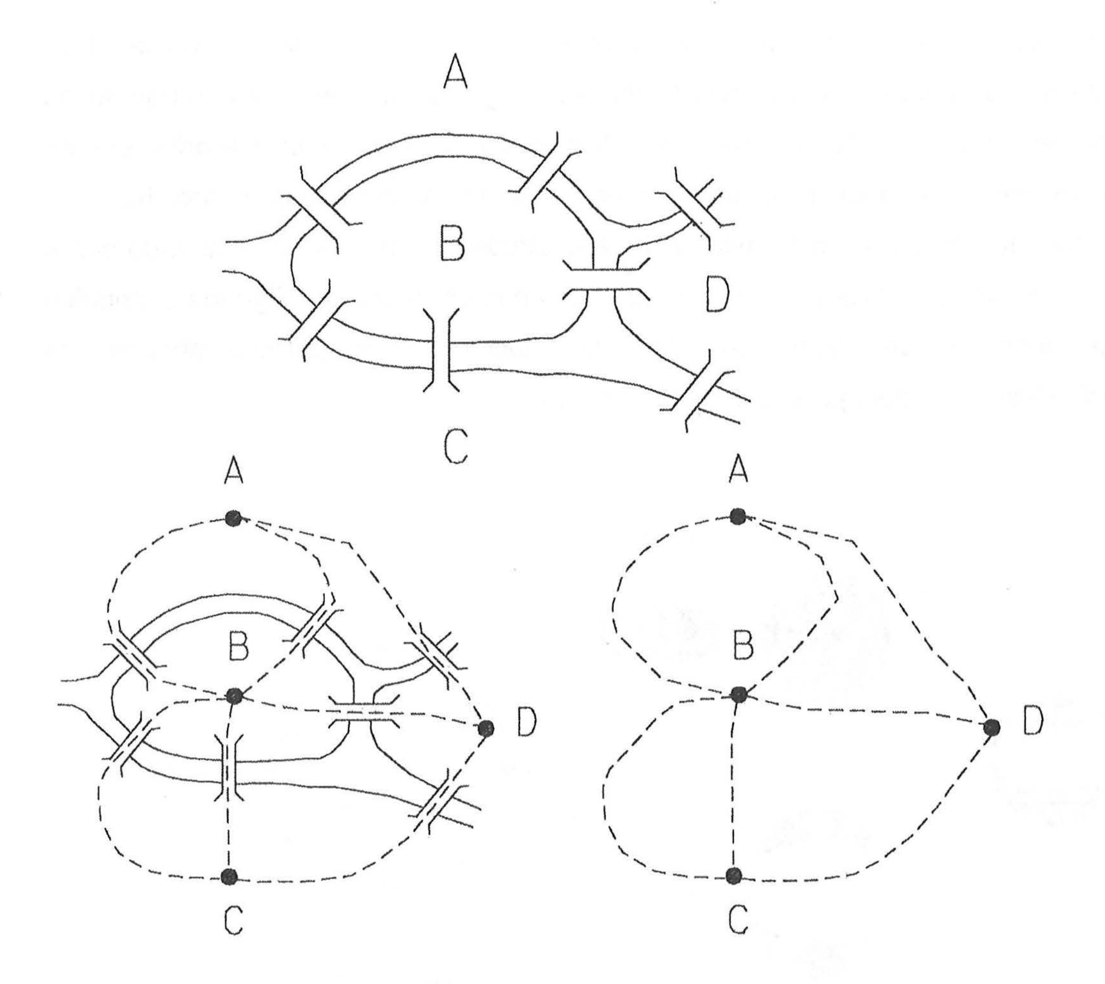
Por otra parte, las celdas pueden estar *orientadas*, es decir, que pueden tener un sentido que se indica mediante una flecha a lo largo de cada arco, o con una flecha circular en el interior de los polígonos. Para evitar posibles ambigüedades existen reglas que ordenan la sucesión de los arcos; por ejemplo, tras seguir un arco hasta un nodo, hay que moverse en sentido antihorario alrededor del nodo, continuando por el primer arco hacia el exterior. De este modo los puntos, arcos y polígonos orientados proporcionan un entramado completo de relaciones que permite analizar sin ambigüedades la topología existente en el plano.



D/ Espacios en red.

Se considera a Euler el fundador del estudio sistemático de la topología, y ello se ha asociado al conocido *Problema del Puente de Königsberg sobre el río Pregel*, que requería trazar un circuito por los siete puentes que unían ambas orillas y dos islas entre sí, pasando sólo una vez por cada uno y vollviendo al punto de partida.

Euler demostró que era imposible construyendo un modelo espacial que hacía abstracción de todo menos de las relaciones topológicas que existían entre los puentes, y concluyó que para conseguir tal circuito el número de lados incidentes en cada nodo intermedio había de ser par.

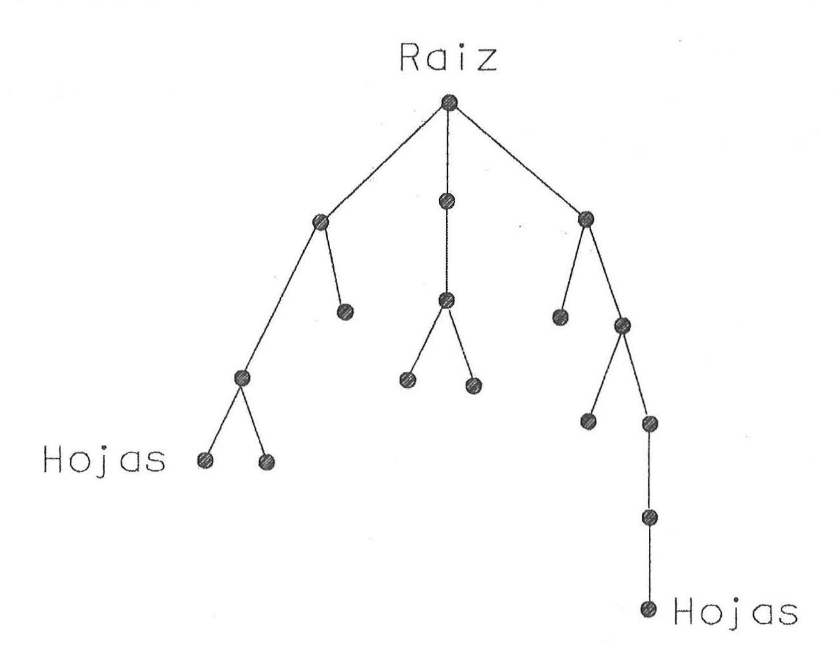


Un *grafo* es un conjunto no vacío de nodos con un conjunto no ordenado de pares de lados incidentes en aquéllos. Sólo representa la conectividad existente entre elementos del espacio, y proporciona un modelo muy abstracto de relaciones espaciales.

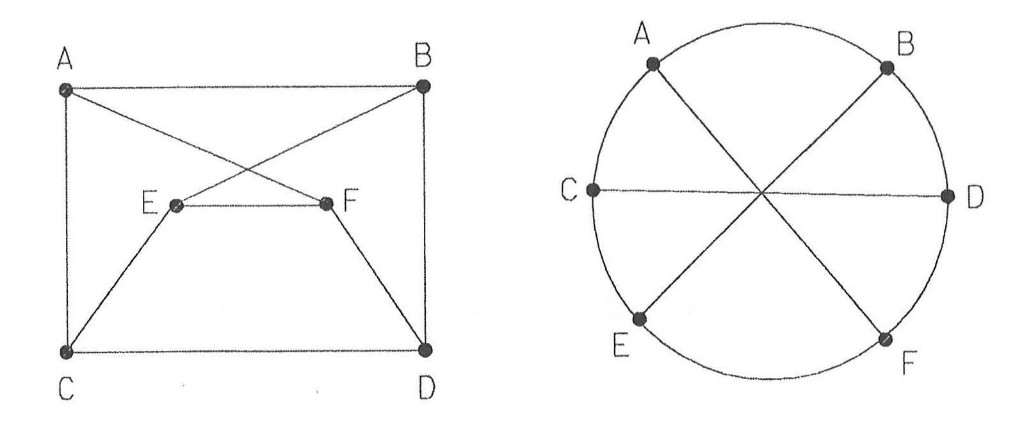
Un grafo puede ser:

- *Dirigido*, cuando cada uno de sus lados tiene asignada una dirección (simbolizada por una flecha).
- *Etiquetado,* cuando cada lado tiene asiganad una etiqueta (número o cadena de caracteres), que se indica en un diagrama próximo a los correspondientes lados.
- *Cíclico,* cuando la dirección va desde un nodo hacia sí mismo, atravesando al menos un nodo.

- *Árbol,* que es un grafo conexo acíclico.
- *Árbol con raíz*, es un árbol que posee un nodo diferenciado, habitualmente representado en la parte superior, con los inmediatos *descendentes* y *hojas* terminales.



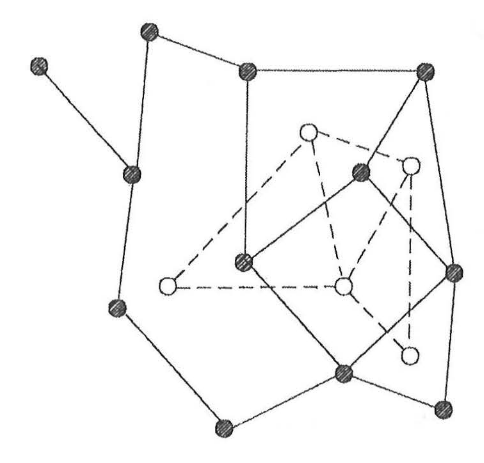
Además, un grafo acíclico dirigido define un orden en los nodos; es decir, que como hemos visto anteriormente, define una relación transitiva, reflexiva y antisimétrica en un conjunto.



Por otra parte, dos grafos son *isomorfos* si muestran las mismas relaciones de conectividad.

Un concepto muy útil en los SIG asociado a los grafos planos es el de la *dualidad*: el *dual* G* de un grafo plano G se obtiene asociando a cada cara de G un nodo en G*. Dos nodos en G* estarán conectados por un lado si, y sólo si, sus cara correspondientes en G son adyacentes.

Análogamente, dado un lado *l* en G, el lado dual *l** de *l* une los nodos en G* correspondientes a las dos caras en G incidentes en *l*.



En los SIG, este concepto se utiliza en casos especiales de representación de objetos geométricos cuales son las superficies sin contornos -representación de datos *quad-edge*.

E/ Los Espacios Métricos.

Se consideran *métricos* aquéllos espacios en los que se ha incorporado el concepto de *distancia*.

Es decir, que un conjunto de puntos A es un espacio métrico si existe una correspondencia entre ellos y los números reales -llamados sus distancias-, de modo que:

- Para cada par de puntos p_1 *y* p_2 en A, *distancia* (p_1, p_2) > 0, si p_1 *y* p_2 son puntos diferentes, y *distancia* (p_1, p_2) = 0 si p_1 *y* p_2 son idénticos (axioma de separación).
- Para cada par de puntos p_1 *y* p_2 en A, la distancia entre p_1 *y* p_2 es igual a la que existe entre p_2 *y* p_1: *distancia* (p_1, p_2) = *distancia* (p_2, p_1) (simetría).
- Para cada tres puntos cualesquiera en A p_1, p_2 y p_3 , la suma de distancias entre p_1 y p_2, y entre p_2 y p_3, siempre es al menos igual a la distancia existente entre p_1 y p_3: *distancia* (p_1, p_2) + *distancia* (p_2, p_3) . *distancia* (p_1, p_3) (desigualdad del triángulo).

Si las propiedades topológicas proporcionan la posición relativa de las entidades espaciales -rodeado por, dividido por, lindante con...-, las funciones de distancia permiten medir las magnitudes espaciales y proporcionan una posición precisa de los objetos espaciales. Se podría decir que los datos métricos suponen un paso más en la definición de tales objetos.

Los espacios geográfico y cartográfico son a la vez topológicos y métricos, porque todo espacio métrico es topológico[4].

En general, los entornos de trabajo de los SIG adoptan una geometría bidimensional, por lo que la altitud se maneja como un atributo de ciertos puntos del plano.

Por otra parte, una de las formas de representación del espacio geográfico en los SIG consiste en su definición como una colección finita de celdas contiguas y disjuntas, que no permiten su descomposición en partes de dimensiones menores, es decir, en *mapas discretos.* En ellos, están implícitas las entidades geográficas como listas de las celdas que los constituyen -o explícitamente como una estructura en árbol o en *quad-tree*, véanse las estructuras raster-, y que al "pegarlas" tienen unas dimensiones equivalentes a las del conjunto correspondiente.
Evidentemente el modelo raster asume un espacio métrico discreto, que permite situar los objetos en el espacio y medir magnitudes espaciales. En definitiva, asocia a cada

[4] La métrica define el concepto de círculo abierto $[B(x,r) = \{y \in E \mid d(x,y < r\}]$, y a partir de él, el concepto de entorno.

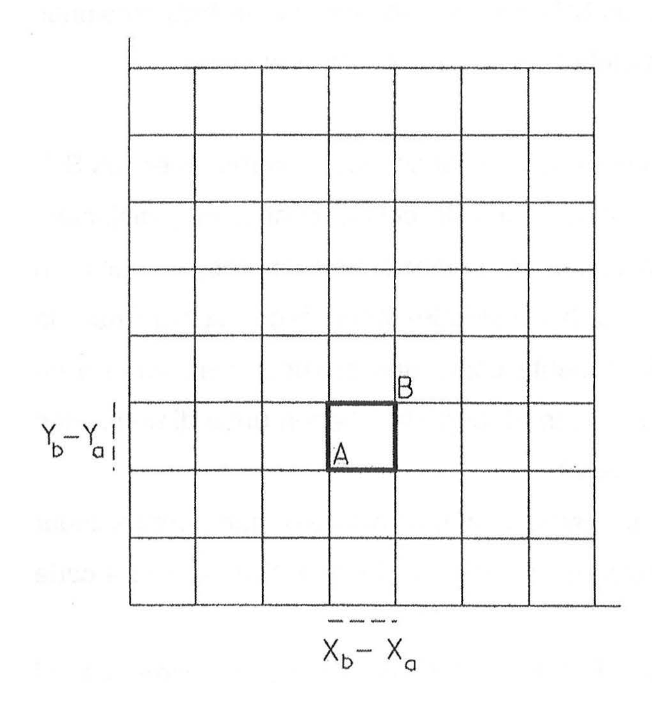
celda un par de números enteros que considerarse como sus "coordenadas", y proporcionan su posición absoluta en el plano, con una precisión que depende de los parámetros de la trama subyacente -cuanto mayor sea la celda, menor será la resolución.

Por otra parte, el modelo raster no podría representar objetos sin dimensión física (puntos) o de una dimensión (segmentos rectos), ya que las celdas son bidimensionales; por ello, para que el modelo sea factible hay que sumir que estos objetos se expanden hasta ocupar la totalidad de la celda (celdas en el caso de los datos lineales) en la que se encuentran.

Por tanto, para medir magnitudes como la longitud se traslada el problema del conjunto de los pares de números enteros $\square^2$ al conjunto de los pares de números reales R^2 -que son la base del modelo vectorial-, y se efectuan los correspondientes cálculos aritméticos.

No obstante, si se mantiene la aproximación discreta se puede utilizar la *distancia de Manhattan* -cuyo nombre proviene de su trama urbana ortogonal-, que mide la distancia por diferencias entre las posiciones medidas en dirección del eje horizontal -o de las filas-, y las medidas en dirección del eje vertical -o de las columnas-:

$$M_{ab} = |X_b - X_a| + |Y_b - Y_a|$$



En cambio, si operamos con pares -o ternas en el espacio tridimensional- de números reales e imaginamos dos ciudades A y B, algunas de las funciones de distancia que un SIG podría aplicar sobre ellas, consideradas objetos puntuales a una escala pequeña, son:

- La *distancia euclídea*, que es la distancia en línea recta que se deduce mediante la aplicación del Teorema de Pitágoras:

  $$D_{ab}^2 = (X_b - X_a)^2 + (Y_b - Y_a)^2$$

- La *distancia geodésica*, que es la distancia medida a lo largo del círculo máximo que pasa por los centros de ambas ciudades.
- La *distancia esférica de Manhattan*, que es la diferencia entre sus latitudes y longitudes respectivas.

Sin embargo, en los SIG se utilizan también otros conceptos de distancia vinculados a la experiencia real extraida del dominio de aplicación; entre éstos se encuentran:

- El *tiempo de viaje*, que es la duración mínima del viaje entre ambas ciudades, utilizando una secuencia de medios de transporte conectados. En este caso concreto, la función no tiene por qué ser simétrica, ya que se puede tardar más en recorrer el trayecto en un sentido que en el inverso (debido a factores como el viento, por ejemplo). Esto implica que el espacio en el que se aplica este concepto de distancia no es métrico.

  En la práctica resulta frecuente considerarla simétrica, aunque hay que tener en cuenta que ello repercute en los resultados de los análisis espaciales.
- La *distancia efectiva*, que se ve influida por factores que intervienen en la percepción de la misma, y que hacen que pueda parecer más o menos larga. Es una medida subjetiva que hay que manejar con precaución y que tampoco produce espacios métricos.

El espacio métrico más utilizado en los SIG es el espacio eulídeo, que además y como ya se ha visto posee una topología "natural". Pero que sea la elección más habitual, no quiere decir que sea la única posible, y en cada caso habrá que estudiar el tipo de espacio que más se adapte a la resolución de los problemas concretos que se desea modelar y analizar.

G/ Geometría de la superficie esférica.

En los Sistemas de Información Globales que trabajan abarcando todo el planeta, es necesario considerar el geoide en su totalidad. Pero como en el espacio euclídeo éste es topológicamente equivalente a una esfera, se pueden aplicar las propiedades de la superficie esférica.
Por otra parte, como ésta es bidimensional, todo punto situado sobre ella quedará definido por dos números -por ejemplo, la longitud y la latitud-, que asumen que la esfera está integrada en un espacio euclídeo tridimensional, cuyo origen del sistema de ejes de referencia coincide con el centro de la esfera. A partir de los ángulos que forma con los ejes el segmento que une el punto con el origen de coordenadas, se define la posición de éste.

La topología de la esfera es similar a la del plano euclídeo, pero plantea diferencias significativas porque no puede haber parte del espacio no cercada, lo que obliga, entre otras cosas, a modificar el Teorema de Jordan.

Algunos tipos de distancia, como la geodésica o la esférica de Manhattan resultan de gran utilidad en este tipo de geometría.

En resumen, hasta aquí se ha efectuado un repaso a las posibilidades de formular matemáticamente el dominio de aplicación, es decir, los diferentes elementos que existen en el espacio geográfico que se estudia, y las relaciones que entre ellos existen. Por tanto, se han sentado las bases para construir, con los datos geográficos disponibles, el modelo de espacio más adecuado a los fines de cada SIG concreto.

Comentarios bibliográficos

Sobre la superposición de datos procedentes de fuentes diversas, es interesante leer el método seguido por MCHarg (1969).
Burrough (1986) expone muy claramente los errores más frecuentes que aparecen en los SIG, aunque con un claro enfoque hacia los utilizados en las Ciencias de la Tierra.

Sobre la historia de la evolución de los SIG se recomienda la lectura de los tres artículos de Steinitz (1993a, b y c), uno de los precursores de la tecnología SIG, aparecidos en la revista *GIS Europe*.

Sobre los conceptos matemáticos que se exponen en el texto, se puede consultar para el tema de las transformaciones y curvas de Jordan, el tomo I de Puig Adam (1986), y sobre proyectividad, el tomo II.
Sobre Topología y Espacios Métricos, conviene consultar el capítulo X de Arnaudières y Fraysse (1988).
Para profundizar en la Teoría de Conjuntos, el de García García y López Pellicer (1984), y el Apéndice 1 de Villa (1994).

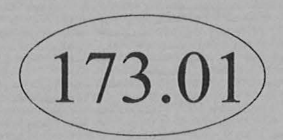
Finalmente, Cebrián (1994) expone con claridad en los capítulos I y II los conceptos de los espacios topológicos y métricos en los SIG. También trata Worboys (1995) el tema en su capítulo 3, aunque conviene revisar algunas de sus definiciones matemáticas.
Si se desea tener una aproximación al diseño de los modelos, recomendamos el completo ejemplo que desarrollan Salas y Chuvieco (1993) para variables microclimáticas, que aportan además una interesante bibliografía
También sobre el desarrollo de modelos de aplicación en Arqueología y Antropología se recomienda la consulta de Aldenderfer y Maschner (1996), que incluye en las páginas 190-201 la experiencia del SIG arqueológico de la Comunidad de Madrid elaborado por un equipo de la Universidad Autónoma.
Y en las actas de los congresos de AESIG se pueden encontrar también experiencias bien explicadas como la que desarrolla el *SinambA* en la comunicación de Moreira y Azcárate (1993), u otras como las expuestas por Casas (1993), por García y Medina (1993) y por Castro y García-Abad (1993).

NOTAS

NOTAS

NOTAS

CUADERNO



CATÁLOGO Y PEDIDOS EN

http://www.aq.upm.es/of/jherrera

info@mairea-libros.com

I.S.B.N. 84-9728-116-0

84-9728-116-0